TABLE OF CONTENTS

BIOLOGY AND BIOTECHNOLOGY RESEARCH PROGRAM
February 2004

Our Mission—Strategic bioscience for national security, health and the environment

LLNL conducts multidisciplinary bioscience to fill national needs. Our primary objectives are to:

- develop knowledge and tools for biodefense and related infectious disease prevention, diagnosis and treatment
- understand and utilize biological processes relevant to the environment and energy production
- understand and mitigate health effects of radiation and chemical exposure
- develop biochemical measurement and computational modeling capabilities which enable understanding of biological processes
- develop technology and tools which enhance healthcare

We execute our mission through core competencies in:

- microbial and mammalian genomics - the characterization of DNA, the genes it encodes, their regulation and function and their role in living systems
- protein function and biochemistry - the structure, function, and interaction of proteins and other molecules involved in the integrated biochemical function of the processes of life
- computational modeling and understanding of biochemical systems - the application of high-speed computing technology to simulate and visualize complex, integrated biological processes
- bioinformatics - databasing, networking, and analysis of biological data
- bioinstrumentation - the application of physical and engineering technologies to novel biological and biochemical measurements, laboratory automation, medical device development, and healthcare technologies

Our role within the broader Laboratory is to:

- Solve challenges of national significance
- Develop a research base which prepares the Laboratory for new missions and needs
- Provide an environment which attracts and develops strong bioscientists
- Serve as a conduit for applying LLNL capabilities Laboratory-wide to biological challenges
- Maintain an infrastructure of facilities, instrumentation, people and legally required oversight.

We utilize the Laboratory's exceptional capabilities in the physical, computational, chemical, environmental and engineering sciences.

We partner with industry and universities to utilize their state-of-the art technology and science and to make our capabilities and discoveries available to the broader research community.

1.0 Research Areas and Achievements

Research Areas

The Biology and Biotechnology Research Program includes six elements that reflect the research areas of interest for the Program and reflect the scientific thrusts that comprise our strategic plan for bioscience at LLNL.

Biodefense Division
Our work in the biodefense field focuses on providing both the basic bioscience and the tools necessary to render bioterrorism ineffective. Our work focuses on such diverse topics as detection of biowarfare threats, human and microbial forensics research and applications, and presymptomatic disease detection. We are building advanced detection systems to provide early warning, identify populations at risk and contaminated areas, and facilitate prompt treatment. We develop DNA signatures and biological forensics technologies to identify an infectious agent, its geographical origin, and/or the initial source of infection. Similar approaches are applied to human forensics, and are used in both law enforcement and intelligence-gathering activities. To carry out this work, our division has a strong partnership with the Nonproliferation, Arms Control, and International Security Directorate (NAI) here at LLNL, and is supported by a wide range of internal and external funding sources.

Computational and Systems Biology Division
The goal of the Computational and Systems Biology Division (CSBD) is to be a nationally recognized research program in computational and systems biology that is built upon LLNL strengths in computations and simulation science, with strong internal and external collaborations. The CSBD currently consists of four research groups with a total of approximately 20 staff. The division uses a wide range of chemical modeling methods, ranging from homology-based protein structure prediction to highly computationally intensive first principles molecular dynamics. In addition to the computer modeling groups, the CSBD includes an experimental protein science group that is working on several joint experiment/modeling projects. The CSBD's science is highly collaborative, with more than a dozen collaborations with BBRP and external research groups. The division is currently funded by a combination of NIH, DOE, and LLNL LDRD grants.

Genome Biology Division
Research in the Genome Biology Division is focused on generation of genomic resources and on developing tools and strategies for turning these basic data into information regarding genome evolution, gene regulation and biological function. The Division includes LLNL's members of the DOE Joint Genome Institute's (JGI) comparative mapping, DNA sequencing and annotation teams, and the I.M.A.G.E. Consortium, a

central element of the international effort to document expressed sequences in a wide variety of eukaryotic genomes. Genome Biology also includes research groups using comparative genomics to identify critical DNA sequence elements in complex genomes and to trace conservation and change of these sequences in vertebrate evolution. Members are working specifically to (1) develop new computational tools for genome comparison and characterization of regulatory sequences and modules associated with co-regulated genes; (2) test and document the functions of predicted genes and regulatory elements *in vitro* and *in vivo* using model systems including mouse, chicken and the frog, *Xenopus tropicalis;* (3) track the biological impact of lineage-specific genes with special focus on those encoding zinc-finger transcription factors, the rapid evolution of which may play a central role in modeling and remodeling vertebrate regulatory networks; (4) unravel basic mechanisms involved in gene repression, including those underlying genomic imprinting; and (5) development of a unique set of mouse models for studying human gene function and inherited disease. Our goal is to provide genomic data, novel computational and experimental tools, and information regarding functions, regulation and evolution of vertebrate genes as a resource to the biological research community.

Health Effects Genetics Division

The long-standing research objectives for the scientists of the HEG Division have been to understand the mechanisms of radiation and chemical toxicities in cells so that we can reduce the individual and population health consequences of exposures through improved detection, prevention, and intervention. There is a major emphasis to identify and characterize the biochemical mechanisms that detect DNA and other forms of cellular damage and those physiological and genetic mechanisms that determine cell survival, genomic integrity, and tissue toxicity. Research themes include mechanisms of DNA repair, food mutagenesis, chemical carcinogenesis, genetic susceptibility, developmental toxicity, and radiation biology. Recent HEG research has lead to advances in understanding DNA repair, reducing the mutagenic effects in food preparation, detecting genetically defective somatic and germinal cells, and interrogating the genome-wide responses of cells exposed to toxicants. HEG research has benefited from the broad expertise provided by multidisciplinary teams, the availability of advanced biocomputing, and the implementation of enabling genomic and proteomic technologies.

Molecular Biophysics Group

The goals for this focus area are to develop and apply new research tools based on LLNL's capabilities in physical, chemical and engineering sciences to enable the quantitative analysis of cells, their components and the biomolecular processes essential for life. A combination of physical techniques and complex simulations and modeling are used to obtain and interpret experimental data. Our current tools, which include x-ray crystallography, nuclear magnetic resonance spectroscopy, mass spectrometry, various forms of spectroscopy (single molecule fluorescence, PIXE, XAFS, Raman) and microscopy (scanning probe, confocal) and several protein over-expression and purification technologies, are applied to the 1) structural and functional analysis of proteins and their complexes, 2) characterization of protein-nucleic acid interactions, 3) development of synthetic ligands that bind selectively to unique sites on the surfaces of proteins, and 4) the

development of new metabolite-based methods for monitoring the functional state of an organism. Our research focuses on a wide variety of efforts ranging from bidefense to cancer, with funding from NIH, NSF, DHS, LDRD and Industry.

Environmental Microbiology Group
This is a new thrust area for BBRP that will bring "biology of scale" to microbial biology and apply this knowledge to the DOE missions in energy, security and environmental protection and restoration. We plan to grow this research area to support the DOE/OBER Genomes to Life Initiative.

Bioinformatics Group
The BBRP Bioinformatics Group consists of scientists, programmers, programming technicians, system adminstration, computer security, and desktop support personnel. The group supports ongoing work within the BBRP program and the Joint Genome Institute in Walnut Creek. Group members are involved in a wide variety of projects, such as high-throughput DNA sequencing, whole-genome analysis and annotation, biostatistics, comparative genomics, I.M.A.G.E. cDNA library analysis and maintenance, mapping, and administrative systems development and support.

Physical Biosciences Institute
The Physical Biosciences Institute (PBI), a new LLNL University Relations Program Institute, officially opened in 2003. The goal of the PBI is to incubate new post-doctoral projects that use LLNL's advanced analytical capabilities to address questions in quantitative biology at the single-cell and single-molecule level. All of these projects involve collaborations with directorates outside the BBRP and most of the projects involve collaborators at a University of California campus.

Interaction between these research areas and related research programs that exist across the Laboratory help to weave a strong Bioscience Program at LLNL.

(this page intentionally left blank)

Research Abstracts

The following abstracts highlight research being reviewed from 2003 by our scientific advisory board. Titles and authors are listed below for abstracts found on the following pages.

Tracking of dual-labeled SiHybrid gene-silencing molecules in individual human cells using confocal laser scanning microscopy
Larry Dugan

Single-cell level investigation of cytoskeletal response to external stimuli
Amy Hiddessen

Single molecule studies of chromatin
Christopher Jeans

Application of SERS nanoparticles for intracellular pH measurements
Ted Laurence

Using femtosecond laser subcellular surgery as a tool to study cell biology
Nan Shen

Force spectroscopy to study multivalent binding in protein-antibody interactions
Todd Sulchek

Development of Genomes to Life Technology

Rod Balhorn and Dennis Matthews

Genomes to Life, a major new program being developed by the Department of Energy as its contribution to the life sciences over the next two decades, focuses on understanding how the genomes of individual and communities of microbes perform the activities needed to survive in harsh environments, process hazardous compounds, and interact to perform the complex functions that enable their survival and help maintain those features of our environment that are essential for the existence of higher forms of life. In an effort to facilitate our achieving this understanding, the DOE has developed a plan to create four facilities over the next decade that function as factories, producing microbial proteins and affinity reagents specific for each protein, characterizing the proteomes of each organism under relevant environmental conditions, churning out basic information about the protein complexes present in each microbe that define how the proteins interact to function, and applying computational and experimental methods to discover the interplay between all the functioning protein complexes that constitute the living organism.

During the last year, LLNL has begun both to develop a strong scientific program in microbial science and to position itself to be a major contributor to the development of the new technology needed to make these four facilities a reality. Working with investigators throughout the LLNL (Chemistry and Materials Science, Engineering, Energy and Environment, Physics, Computations) and key collaborators in industry (Protasis, Somalogic) and at other institutions (UC Davis, UC San Francisco, UC Berkeley, Oak Ridge National Laboratory, Argonne National Laboratory), we have begun identifying specific technologies LLNL could develop and contribute to these protein production and characterization “factories”. These include new methods for isolating, fractionating and manipulating all the proteins in a microbial cell, the high-throughput production of protein-specific affinity reagents, the design of new tags for visualizing proteins and monitoring their function, computational methods for predicting protein structure and function, new experimental methods for identifying protein and complex function, new imaging methods for analyzing individual proteins and protein complexes in vivo, and the development of new methods for rapidly obtaining high resolution structural information from single protein molecules and protein complexes without crystallization.

Comparative Genomics for Biodefense

Patrick Chain

We are applying our resources in genomics and comparative genomics to help better understand microbial pathogens of interest to biodefense. Knowing the sequence of a target organism allows the extraction of information to begin understanding metabolic and pathogenic potential, species variation, differences between species, genome and pathogen evolution, which can be used for the development of methods to detect and differentiate between microbes, and thorough understanding of the underlying mechanisms of virulence to create rational vaccines and treatments. Work undertaken has included a variety of pathogens comprising both viruses and bacteria.

Together with the Centers for Disease Control (CDC) and United States Army Medical Research Institute of Infectious Diseases (USAMRIID), we have completed the genomes of BioSafety Level 4, CDC Category A viruses, including seven Filoviruses (including Ebola and Marburg) and eight Arenaviruses (including Lassa and Sabia). These sequences are being used to develop DNA markers specific to these lethal pathogens for use in detection.

We have completed or nearly completed the genomes of several *Francisella* strains. *F. tularensis* is the least well characterized Category A bacterial pathogen, as very few virulence factors have been reported and in general the mechanism of pathogenicity remains unknown. Comparing the genomes of a highly virulent strain, a less virulent strain, an attenuated strain, and a related *F. philomiragia* strain may help elucidate which genes/functions are involved in increased virulence.

In the area of *Yersinia* biology, we have recently conducted comparative analyses of *Y. pseudotuberculosis* with two available *Y. pestis* strains. Several regions unique to either species have been identified and verified using a panel of *Y. pestis* and *Y. pseudotuberculosis* strains. Many differences, from large genomic rearrangements to single point mutations that affect open reading frames to variation of IS elements and their distributions, have been studied. Global gene expression studies have been performed to identify gene regulation under temperature shift and pH shift conditions. Further work in generating targeted knockout mutants in virulent *Y. pestis* has been performed to study the role of specific genes in pathogenicity in a mouse model.

Proteomic Characterization of Host-Pathogen Interactions

Sandra McCutchen-Maloney

It is clear that pathogens such as *Yersinia pestis*, the causative agent of plague, still pose a significant threat to human health both through natural environmental exposures and through intentional acts of bioterrorism. Efforts in the Biodefense Proteomics Group have focused on proteomic characterization of host response to *Y. pestis* and other human pathogens using cell-based, whole blood and animal model systems. Using advanced proteomic technologies including multiplexed 2D DIGE and SELDI-MS, thousands of host proteins have been detected in these host model systems of which hundreds have been differentially expressed after pathogen exposures. These expression patterns clearly distinguish exposure to three closely related pathogens based solely on proteomic host response. In addition to host response, efforts have also focused on proteomic analysis of pathogens to characterize virulence using gene knockout approaches followed by proteomic and phenotypic characterization using high-throughput array technologies. The ultimate goals of this work are to define pathogen-specific host signatures toward presymptomatic detection of disease, to define novel therapeutic targets for infectious diseases such as plague, to define pathogen signatures for detection of naturally occurring and genetically engineered pathogens, and to better define the mechanisms of virulence and pathogenicity.

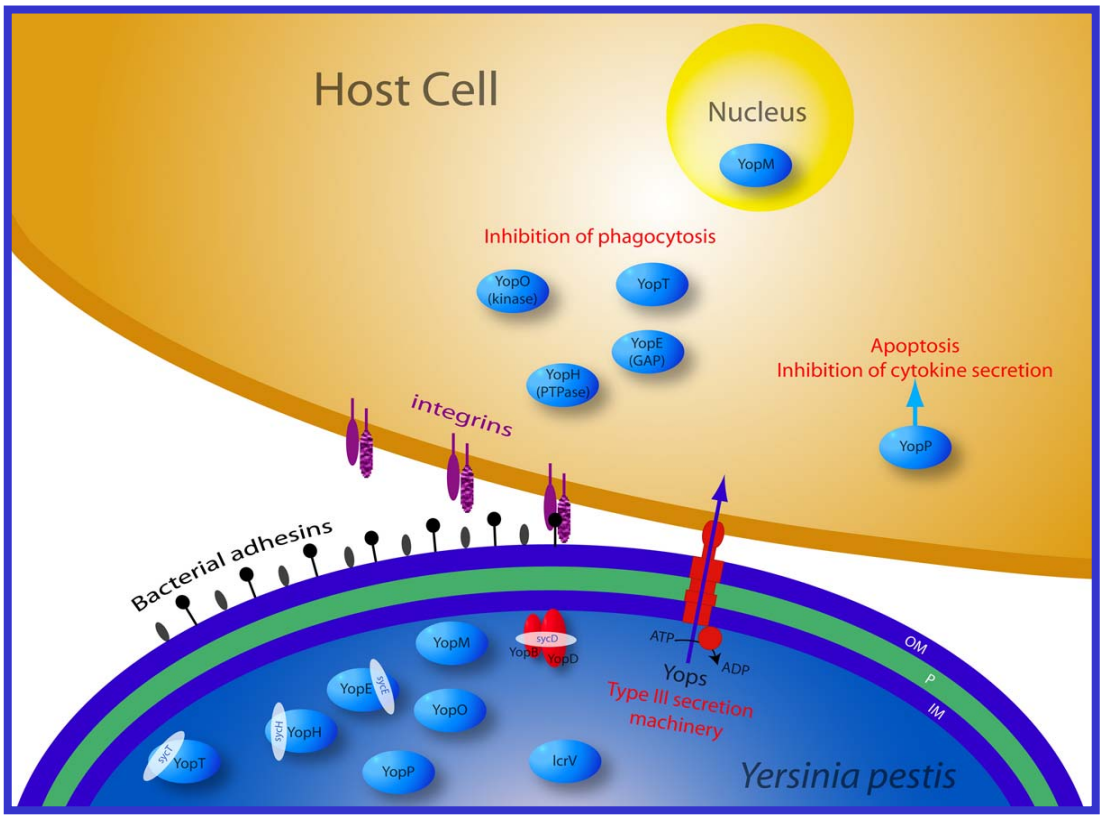

Figure legend: The study of host-pathogen interactions can be used to define the next-generation biomarkers for early detection of biological exposure, with implications for infectious diseases as well as national security.

Pathomics

Richard G. Langlois

"Pathomics" is a comprehensive strategy to develop a systems level understanding of host responses to infectious disease agents, particularly those that represent a threat from bio-terrorism. Its ultimate goal is to predict all possible routes of pathogenicity and host response from our understanding of the genes involved. This will ultimately lead to detection systems to determine that an attack or disease outbreak is beginning prior to people becoming overtly sick. This is critical to our ability to deal with the public health threats from contagious naturally emerging diseases and, importantly for bio-terrorism, genetic manipulation. This is a Manhattan Project-type "Grand Challenge" problem that ultimately will require decades of research. Our current Strategic Initiative Project proposes to establish a new program direction by building the necessary internal multidisciplinary, cross-directorate team; establishing the crucial strategic collaborations with academia and industry; and most importantly focusing on a few specific goals that rigorously proves the paradigm's feasibility.

The initial goals of this Project are to determine if blood samples from humans and animals can be used to elucidate the spectrum of molecular changes that occur in infection, and the time course of these changes. The experimental methodologies that will be utilized for these molecular studies include gene expression arrays, immunoassays, mass spectroscopy, and protein gel electrophoresis. Initially, we will compare smallpox vaccination in humans with cowpox infections in mice. A second target will be characterization of respiratory infections in mice and humans.

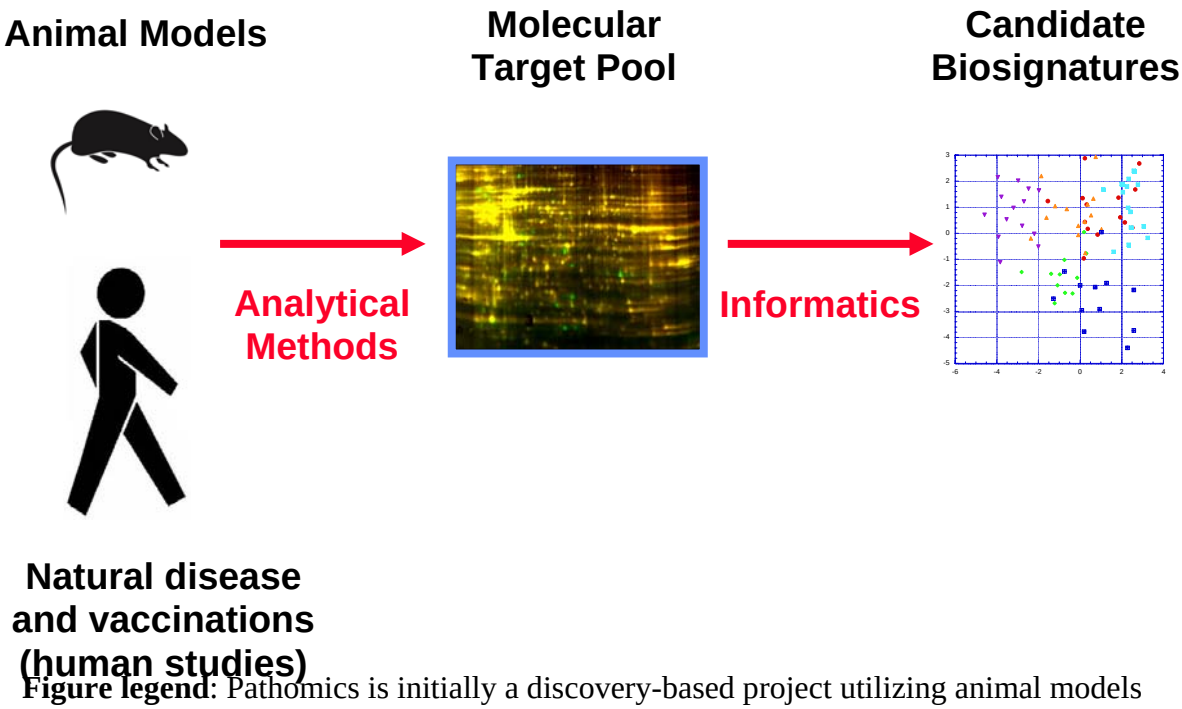

Figure legend: Pathomics is initially a discovery-based project utilizing animal models and natural diseases in humans to identify candidate signatures.

Genome Analysis for Pathogen Detection Signatures

Tom Slezak

Problem Description
Rapid identification of natural and genetically engineered pathogens is vital to our national security. Previously, efforts had concentrated on developing assays that detect organisms by targeting small portions of their genomes (e.g., single, well studied genes). The need arose for new tools that could *rapidly* identify multiple *unique* and specific regions of the genome, increase specificity, and reduce cost of development and time to deployment. This work is a collaboration of computer scientists with BBRP biologists.

Technical Approach/Status
Since August 2000, LLNL has pioneered computational design of DNA signatures for detection of natural disease outbreaks or bioterrorism. Our DNA signature pipeline is a fully automated software system for identifying unique regions on pathogen genomes and selecting optimal sequences for development of real-time detection assays (Figure 1). These signatures are stringently tested by LLNL biologists and by external collaborators. They are validated by the CDC and placed into regular use for environmental monitoring. Air samples are collected on filters, then analyzed for the presence of pathogens (Figure 2). Results can be obtained in as little as 90 minutes. As of early 2004, our assays have over 700,000 live bio-defense uses with zero false-positives.

CY2003 Progress
We scaled our original DNA signature pipeline to run in parallel on a 24-CPU server. We can now process bacterial pathogens (e.g., anthrax) in less than two hours. An invitation from the CDC to provide signature design help led to signatures for smallpox, monkeypox, and SARS--all in response to unplanned world events. Our system now automatically downloads new and updated microbial sequence from multiple public web sites and checks all fielded signatures for potential false positive/negative results. We linked the DNA signature pipeline to a fully automated annotation system, which provides us with the additional capability of down-selecting from among many good signature candidates based on their association with genes of interest (e.g., virulence factors). Leveraging our success with DNA signatures, in 2003 we prototyped a protein signature pipeline that can locate unique protein-based diagnostic targets on a whole-proteome scale. Our system identifies unique peptides, locates them on 3D protein models, and annotates for best-choice selection.

Significance
In 2003, our signatures entered BIOWATCH, a program which daily monitors air in several dozen US cities. We have been asked to apply our computational system to a variety of human, animal, and plant pathogens. Success of our DNA signature design effort has established LLNL at the forefront of pathogen diagnostics. Analyses we performed in 2003 for the CDC have directly influenced the nation’s policy on pathogen defense.

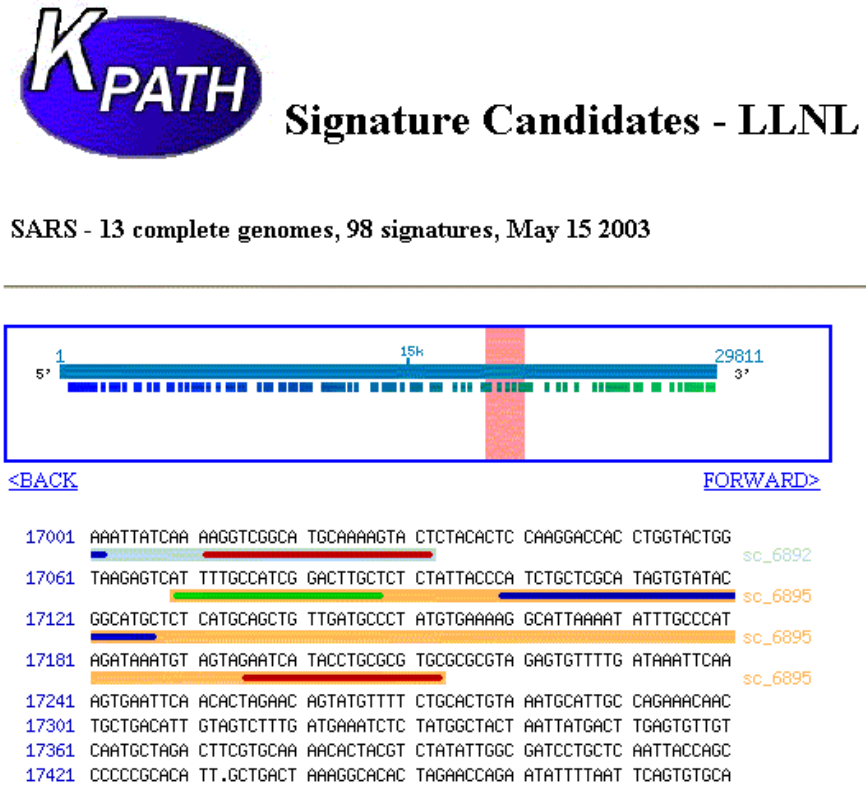

Figure 1. Location of unique sequence signatures in a portion of the genome of the causative organism for SARS as determined by the automated DNA signature pipeline designed at LLNL.

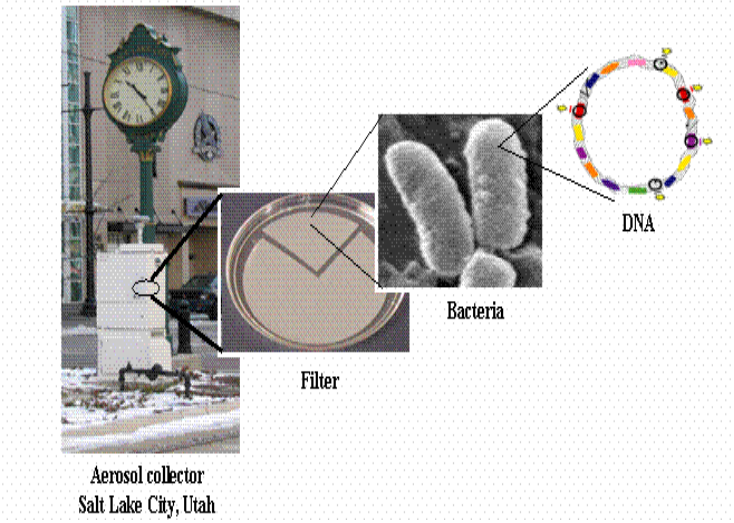

Figure 2. DNA signatures developed at LLNL are used to monitor air for the presence of various pathogenic organisms.

Genetic Assay Development and Utilization

Evan Skowronski

The assay development pipeline at LLNL is designed to deliver multiple independent, highly specific genetic PCR-based assays to detect biological pathogens. The pipeline consists of an *in silico* bioinformatics screening and selection process utilizing known DNA/RNA sequences which generates thousands of potential agent-specific ***signatures***. Each signature is then rigorously screened against multiple strains of target, near-neighbor "look-alike" organisms, and panels of diverse eukaryotic and prokaryotic nucleic acids. Finally, environmental screening against soil and aerosol samples is performed to eliminate, as much as possible, spurious environmental cross reactions. All final signatures are normally screened against a pool of over 2,000 samples, most of which represent complex environmental backgrounds. These signatures are then used to generate ***assays*** that are capable of performing on a wide variety of detection platforms. The resulting assays are then sent to the appropriate agency (CDC, USDA) and validated independently. Many of the validated and widely distributed CDC Laboratory Response Networks nucleic acid assays for CDC Class A pathogens were developed here at Livermore. More recently, a concerted effort to address agricultural pathogens has been made to protect the nation's food supply and protect trade and economic interests. Operational and financial pressures have also driven the development of heavily multiplexed assays to generate a maximum amount of information quickly with a fraction of the cost of a single assay. Among the relevant pathogens that LLNL has been asked to provide candidate signatures for in the past two years include SARS, Monkeypox, West Nile Virus, Exotic Newcastle Disease (END), and Foot and Mouth Disease Virus (FMDV).

LLNL has had the opportunity to respond to real-time infectious disease outbreaks using some of the technologies described above. In October 2002, the presence of Exotic Newcastle Disease was confirmed in a backyard chicken flock in southern California. After providing potential assay targets to the UC Davis California Animal Health and Food Safety (CAHFS) Laboratory, END broke out into the commercial poultry population, triggering a wave of import bans and quarantines in an attempt to limit the spread of the disease. LLNL personnel worked closely with those at UC Davis/CAHFS to optimize the assays for the virulent and vaccine strain viruses, to transition to the processing of thousands of samples per day (a calculated increase in throughput of 50X), and to develop methods to handle "dirty" environmental samples for rapid screening (eliminating the need for a 48 hour viral culture). By August 2002, more than 80,000 samples had been processed, and the state was declared END-free (*Shikora et al*, manuscript in press). Notable in this effort was the ability of the vaccine strain assay to differentiate between virulent and avirulent strains of END, which saved four flocks and >150,000 chickens from depopulation (a > one million dollar indemnity).

BIOWATCH

William Beckwith

The Lawrence Livermore National Laboratory has a well-documented history of developing new technologies to meet the nation's bio-security and microbial detection needs. The national need to quickly and accurately identify bio-threat agents has increased dramatically since the events of September 11th, 2001, and LLNL maintains its position at the cutting edge by constantly demonstrating our evolving expertise in the development and application of state-of-the-art real-time diagnostic methodologies.

The BIOWATCH program is just one example of a refined product with roots that can be traced back to an LLNL-initiated demonstration project. Modeled after LLNL's Biological Aerosol Sentry and Information System (BASIS), the BIOWATCH program began with a presidential directive in January 2003 as a deployment of the Department of Homeland Security. Originally tasked with monitoring urban environments for the presence of numerous bioterrorist threat agents, the program has since been successfully implemented in most of our nation's major metropolitan areas. To date, the entire program has tested over 700,000 environmental specimens.

The diagnostic program itself employs a panel of real-time polymerase chain reaction (PCR) diagnostics all based on the TAQMAN ™ chemistry. Briefly, the bioinformatics/ genome analysis group at LLNL performed an exhaustive bioinformatics analysis, and ultimately determined a panel of candidate genetic signatures to uniquely identify each pathogen's genetic material. Candidate signatures were subsequently screened in the laboratory at LLNL to rule out the recognition of closely related species, and the entire package was delivered as a multi-tiered screening / confirmation panel that, to date, has never yielded a false positive. Continued progress in the advancement of the diagnostic platform will increase the high throughput capabilities of BIOWATCH laboratories and also enhance their capacities to provide more surveillance data in less time.

Next-Generation Biowarfare Agent Detection

Allen Christian

Biodetection is divided into several levels; detection to warn, to protect, to contain, and to treat. Each of these is typically accomplished with different strategies and platforms; LLNL plays a significant role at each level. As examples, the Biological Aerosolized Mass Spectrometer (BAMS) detects to warn, and the Autonomous Pathogen Detection System (APDS) detects to protect. Next-generation detection systems, such as the Biobriefcase (shown below), are currently being developed at LLNL, with the goal of decreasing size and cost, while maintaining effective and reliable area monitoring. The primary function of the Biobriefcase instrument is to continuously monitor the environment for aerosolized biological pathogens. Biobriefcase will be an automated compact device capable of inexpensively performing multiplex assays with excellent sensitivity and specificity. The focus of this project will be to develop and test the platform while miniaturizing the fluidics and detection components resulting in a more compact and cheaper unit to operate. The system itself is designed to be modular, balancing various capability levels with cost, depending on the needs of the end users.

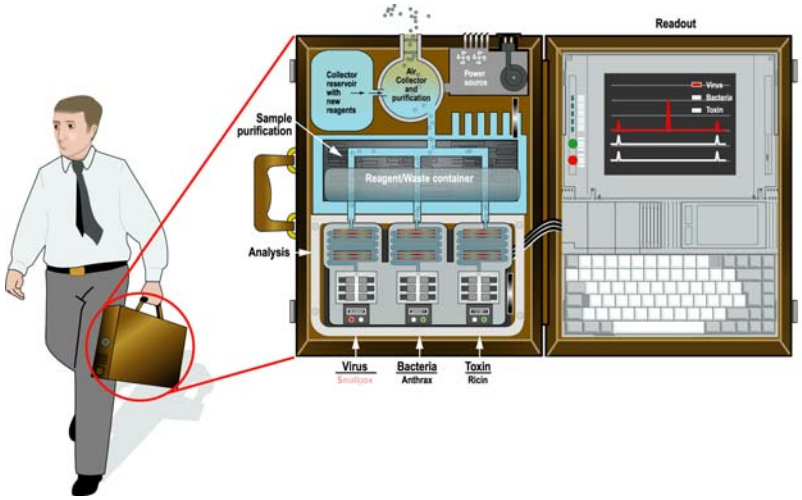

Figure legend: The Biobriefcase is being developed as a portable unit used to monitor the environment for the presence of pathogens.

Protein Model Database

Krzysztof Fidelis, Alexei Adjoubei, George B. Magklaras,
Andriy Kryshtafovych, Pawel Daniluk

Protein 3D structure is a key element of the gene sequence - gene function information spectrum. However, even with rapid advances in crystallography and NMR spectroscopy, and a substantial investment in structural genomics, our knowledge of protein structures lags far behind our knowledge of sequences. Computational modeling techniques offer one of the most attractive means to close this two-order-of-magnitude gap. However, computationally derived models, much more than experimental structures, require that a reliable model quality estimate is presented alongside the model structure itself. Also, any genomic-scale information is especially valuable when considered as part of an easily-navigable knowledge system, allowing for multiple comparisons, and providing access to other, related types of data. With the ultimate goal of creating a public facility, we have set out to address the following aspects of such a protein structure-centered system: (I) evaluation and annotation of model quality; (II) access to the best modeling techniques presently available; and (III) initial development of the technical framework dedicated to comparative structural genomics.

I. The Protein Model Database (PMD) will provide an easy, structured access to several types of information relating to model quality. (1) A clear description of how the model was generated, including modeling group, applied methodology, and the protein-specific details; (2) model quality annotation based on methods' past performance and on the complexity of the modeling task, including expected conservation of structure; (3) model assessment based on stereochemistry and force-field type calculations; (4) consensus analysis in cases where multiple models are available; and (5) whenever possible, annotations by specialists.

II. Access to the best modeling techniques will be sought, including: (1) automated modeling applied to genome-scale structure prediction; and (2) supervised modeling, involving concerted efforts of multiple research groups worldwide, applied to few, high-impact modeling targets.

III. The model database will be run by a relational management system and accessed with specialized search engines over the WWW. The database dictionary that we develop contains definitions of data items in the STAR/mmCIF format. The PMD dictionary describes the many aspects of the protein structure modeling process and the modeled structures. Categories providing direct compatibility with protein sequence, metabolic pathway and protein structure (Protein Data Bank) databases have also been developed. Links with other bioinformatics databases will be established, including comparison of models with experimental structures (PDB), and with genome and protein sequences (e.g. GenBank). Custom search capabilities will be added to include specialized tools allowing for model comparison and visualization.

The increased ability to generate models of protein structure will be demonstrated by application to Laboratory-based projects. The basic nature and general applicability of protein structure data make this project relevant to many Laboratory and DOE core missions, from human health to homeland defense.

Impact of other foods on the risk of exposure to carcinogenic heterocyclic amines in the diet

Kristen S. Kulp, Mark G. Knize, Susan L. Fortson, Jennifer L. Montgomery, Brian J. Bennion, Felice C. Lightstone, Cynthia P. Salmon and James S. Felton

The human diet is a complex mixture of substances that not only provide nutritional sustenance but may also play a role in the causation, modulation or prevention of human cancer. Because diet is one aspect of an individual's lifestyle that may be practically modified, it is important to quantify dietary exposures to understand an individual's risk of cancer and identify dietary interactions that may modulate the risk. High temperature cooking of protein rich foods results in the formation of a group of structurally related heterocyclic aromatic amines (HAs) that are potent genotoxic carcinogens. These compounds have been shown to cause breast, prostate and colon tumors in animal studies. We are investigating the effects of exposure to these compounds using human, cell, and analytical assay systems. Our results demonstrate that these compounds are being absorbed and excreted by people that consume over-cooked meat, that dietary interactions may affect both the levels of exposure and the outcome of the exposure and that these compounds may be having an effect at the cellular level that goes beyond genotoxicity.

To investigate human metabolism of HAs at dietary doses we devised an assay to quantify metabolites in human urine following a single exposure of well-cooked meat. Using this method, we determined that individuals vary in their absorption and metabolism of these compounds and that lifestyle and diet can alter metabolism, which may affect an individual's susceptibility to cancer development. We also investigated the effect of two specific foods, broccoli and soy, that are believed to be preventative for cancer. We found that adding these foods to the diet may affect the activity of metabolizing enzymes and accelerate metabolite excretion.

An important component of exposure to food carcinogens is digestion of the compounds from the cooked food matrix. We developed an *in vitro* digestion model to understand how variation in digestion parameters and the consumption of other foods with the cooked meat can affect carcinogen absorption. Using this model, we found that the digestive enzyme pancreatin most affects the release of HAs from the cooked food matrix and that dietary fiber co-digested with meat binds HAs and decreases their absorption.

To understand why specific organs are targets for HA tumorigenesis, we investigated the effects of PhIP, the most commonly occurring HA in cooked meat, at the cellular level. We found that in human breast cancer cells, PhIP increases cell number 20-30% greater than control and increases activity of the estrogen receptor up to 1.8-fold compared to untreated cells. Three-methyl-PhIP, an isomer of PhIP, is not able to activate the estrogen receptor and when added with increasing amounts of estradiol, appears to prevent estrogen-initiated activation of the receptor. To understand our experimental results, computational docking studies were done with the PhIP isomers and the ligand-binding domain of the human estrogen receptor alpha. The docking studies demonstrate that PhIP compounds can easily fit inside the estrogen-binding cavity and assume several energetically favorable conformations.

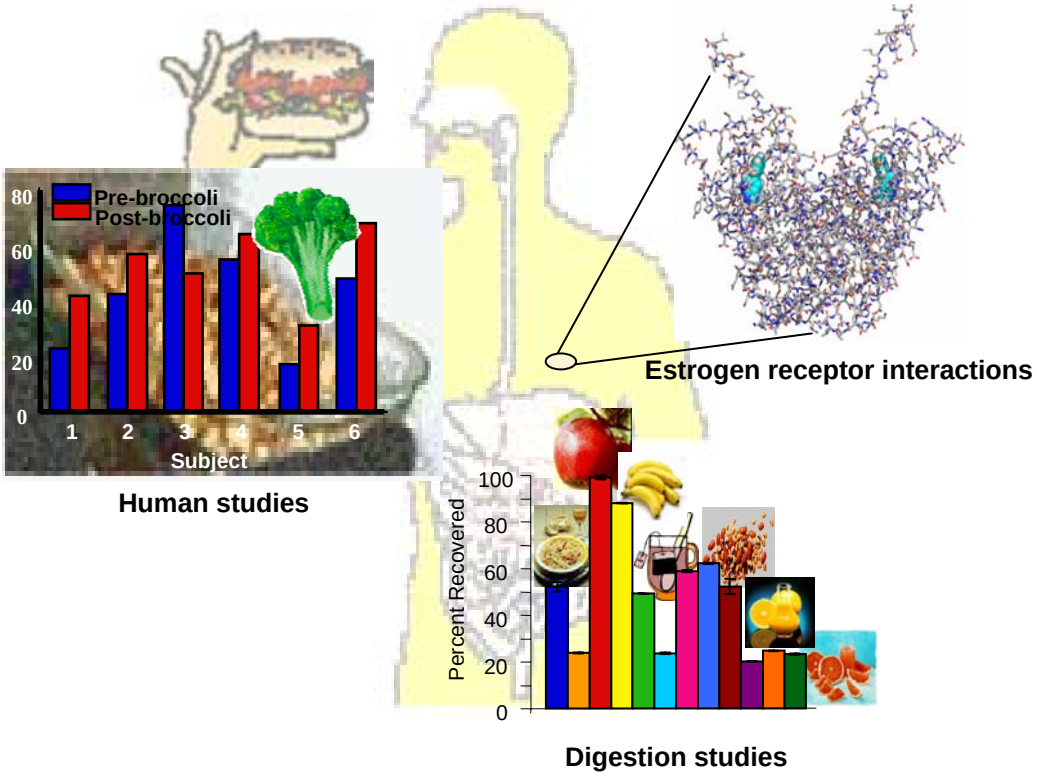

Figure legend: Our studies are designed to investigate the consequences of exposure to food carcinogens on human health. By understanding the factors that determine the absorption, biological activation and cellular effects of these compounds, we can devise potential prevention strategies to reduce the risks associated with HA exposure.

Computational Simulations of Metabolism in Heterocyclic Amines

Felice Lightstone

Heterocyclic aromatic amines are produced in muscle meats and grain products during high-temperature cooking processes such as frying and grilling. A subset of these compounds, the 2-aminoimidazole-azaarene (AIA) mutagens, are characterized by a 2-aminoimidazole ring component. These compounds have been shown to be both mutagenic and carcinogenic; therefore, exposure to these compounds in the diet may increase cancer risk**.**

Mutagenicity is a frequent marker of carcinogenic potential. Identifying the factors that modulate mutagenic potency will give us a better understanding of the human health risks associated with environmental carcinogen exposure. Our goal is to use computer simulations to investigate whether differential cytochrome P450 oxidation is a major factor in determining the mutagenic potency of AIA food mutagens. An important property of these compounds is that they exhibit a 100-million fold range of mutagenic potency while sharing a common, mutagenically active 2-aminoimidazole group.

To study the N-oxidation mechanism of the AIAs in the active site of P450, a homology model of the human P450 CYP1A2 was generated from the available X-ray structure of chimeric rabbit cytochrome P450 2C5. Considering several possible candidates, molecular dynamics was used to refine and then choose an optimum model. The results of the molecular dynamics simulations were compared to known features of the P450 family to identify those most likely to accurately represent those of human CYP1A2. Our CYP 1A2 model demonstrated several relevant P450-specific features, including a kink in the I-helix and structural evidence of a proton shuttle above the I-helix near the heme. Molecular docking studies, using caffeine, PhIP, and 3-methyl-PhIP, were performed on the last MD structure of all the stable CYP1A2 models. One of the CYP1A2 models was able to accommodate caffeine in two conformations. The first conformer is perpendicular to the heme and is the catalytically active conformation. The second conformer is parallel to the heme and is not catalytically active. Interestingly, this binding conformation for caffeine in CYP1A2 was proposed from NMR relaxation experiments.

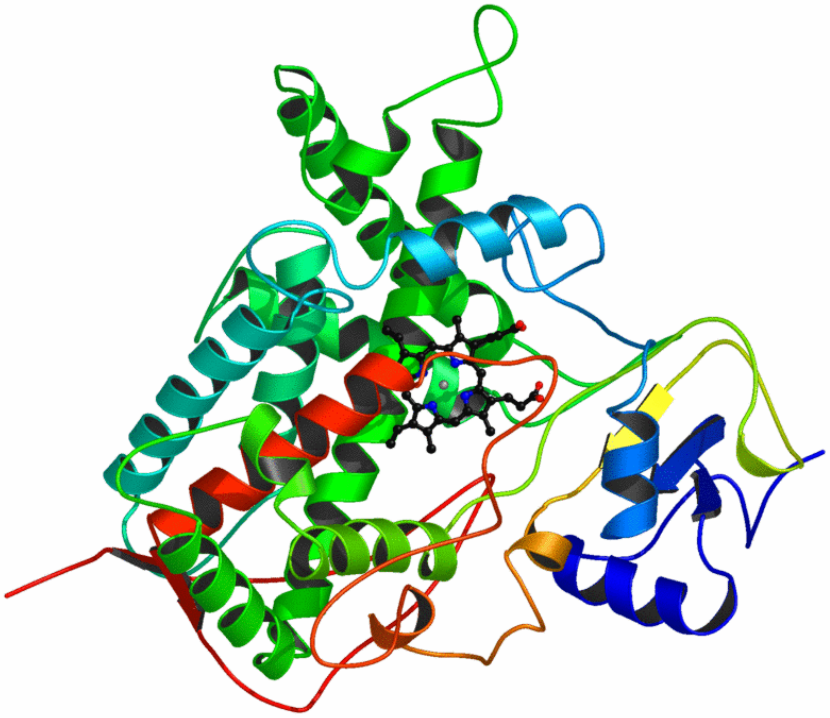

Figure legend: The homology model of the human P450 CYP1A2.

Comparative Genomics

Gaby Loots, Ivan Ovcharenko, Lisa Stubbs

Joint efforts between the Genome Biology Division of BBRP and the Energy, Environment, Biology and Institutional Computing Division are aimed at developing novel biological and computational methods that would provide a framework for decoding the Human Genome. Based on Darwinian evolutionary concepts that functional elements in genomes are under selection, we've developed comparative genomic approaches intended for the identification of evolutionarily conserved elements with potential to code for genes or noncoding regulatory elements. For this purpose we have created a suite of web-based interactive computational tools that allow genomic comparisons for sequences of any length. These include a multiple sequence alignment tool, ***zPicture*** (1) and a multi-vertebrate genome comparative browser, the ***ECR Browser*** (2) that contains pre-aligned sequences for the human, mouse, rat, frog, and three fish genomes (*Fugu*, *Tetraodon* and zebrafish) (see figure). To facilitate the prediction and characterization of putative regulatory elements, we've integrated ***rVista*** (3,4) tool analysis into both the ***zPicture*** and the ***ECR browser*** alignments in order to identify evolutionary conserved transcription factor binding sites in noncoding conserved elements. The ***eShadow*** (5) and the ***Crème*** (6) tools are also designed for characterizing transcriptional regulatory elements. The ***eShadow*** tool performs phylogenetic shadowing in multiple alignments from closely related species and the ***Crème*** tool is designed to identify clusters of transcription factor binding sites in the promoters of co-expressed genes.

These comparative genomic tools are aimed at assisting us to the identify ideal candidate regions to be pursued in functional studies by Dr. Stubbs' and Dr. Loots' groups. Functional characterization of the human chromosome 19 genes, including generation of a catalogue of *in situ* embryonic and adult gene expression patterns, verification and functional examination of transcripts including alternative splicing products and reporter-assay analysis of promoters and enhancer sequences throughout the chromosome, are being performed to annotate and characterize all the genes found on Hsa19. Recently, Dr. Loots has initiated a project that takes advantage of the well-established embryonic manipulations in the developing frog and the whole-genome shotgun sequence that is being generated by the Joint Genome Institute for the frog *Xenopus tropicalis*. By combining comparative genomic strategies, whole mount *in situ* hybridization and transient transgenic techniques in the frog, we plan on identifying and characterizing novel genes shared by all vertebrates which we believe may hold vital roles during the early stages of vertebrate development. Our aim is to couple computational predictions with high throughput *in vitro* and *in vivo* experimentation in all available model systems to develop robust strategies for elucidating the function, regulation, and evolution of vertebrate genes.

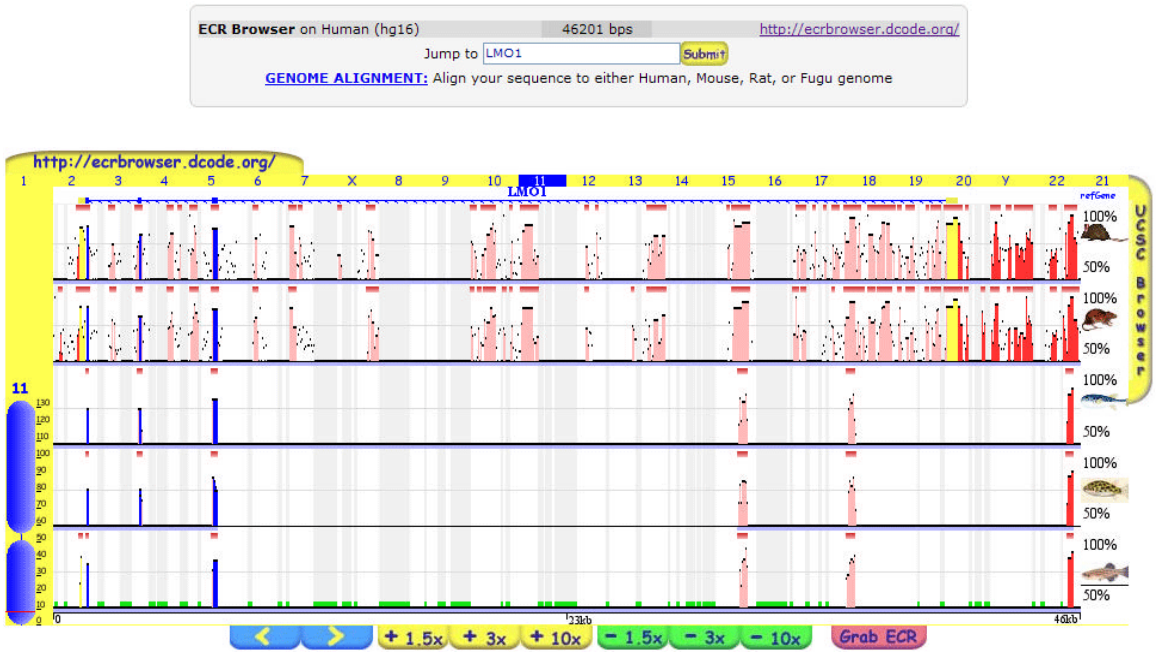

Figure legend: ECR Browser visualization for the Lim Domain Only 1 (LMO1) gene (46kb). The human region (chr11:8209000-8255200; UCSC freeze 16; NCBI Build 34) is compared with the mouse, rat, fugu, tetraodon and zebrafish genomes (the icon to the right indicates the species being compared to human in the pairwise alignment). Each layer contains a pip-type plot that consists of multiple short horizontal black lines. Each line represents an ungapped alignment with the vertical height, where y-axis denotes percent identity and x-axis distance in kb. Exons are in blue, Untranslated regions (UTR) in yellow colors and the direction of the gene is indicated by the arrows. The dark red bar on top of every layer pinpoints evolutionarily conserved regions (ECRs) and are used for color coding conserved elements of different categories: blue - coding exon, yellow – UTR, pink – intron, red – intergenic region.

References;

(1) I Ovcharenko, GG Loots, R Hardison, W Miller and L Stubbs (2004). ***zPicture***: Dynamic alignment and visualization tool for analyzing conservation profiles. *Genome Res* (in press).

(2) I Ovcharenko, MA Nobrega, GG Loots and L Stubbs (2004). ***ECR Browser***: A Tool for Visualizing and Accessing Data from Comparisons of Multiple Vertebrate Genomes. *NAR* (submitted).

(3) GG Loots, I Ovcharenko, L Pachter, I Dubchak and EM Rubin (2002). ***rVista*** for comparative sequence-based discovery of functional transcription factor binding sites. *Genome Res* May;12(5):832-9.

(4) GG Loots and I Ovcharenko (2004). ***rVISTA***: Evolutionary Analysis of Transcription Factor Binding Sites. *NAR* (submitted).

(5) I Ovcharenko, D Boffelli and GG Loots (2004). ***eShadow***: a tool for comparing closely related sequences. *Genome Res* (in press).

(6) R Sharan, I Ovcharenko, A Ben-Hur and RM Karp (2003). ***Crème***: a framework for identifying *cis*-regulatory modules in human-mouse conserved segments. *Bioinformatics*. 2003 Jul;19 Suppl 1:I283-I291.

Tracking of dual-labeled siHybrid gene-silencing molecules in individual human cells using confocal laser scanning microscopy

Larry Dugan

RNA interference is a highly conserved process by which cells degrade unwanted mRNA in a sequence-specific manner using small, interfering RNA molecules (siRNA) and ribonucleases. These siRNA molecules can be made synthetically and induce the same effect when transfected into cells. We have modified the structure of siRNA from double-stranded RNA to RNA:DNA, termed siHybrids or siHYB, and have seen improved gene silencing effects, including increased longevity and level of effect. Fluorescence imaging efforts involving labeling the RNA strand with Cy3 and the DNA strand with Cy5 indicate cytoplasmic uptake of the siHybrid molecules <24 hours after transfection. With this method we could easily detect high levels of labeled siHybrid RNA within the cytoplasm of cells. We have been able to track these molecules in clonal populations of individually transfected cells for up to 10 generations. Furthermore, the use of transfection reagent has been shown to be unnecessary to induce a specific gene silencing effect, which is not the case with siRNA. We are using a number of imaging and analytical techniques to image, track and quantify down to the sub-cellular level the uptake and persistence of siHYB molecules in human cells in order to help elucidate their mechanism of action. We have now shown for the first time a quantifiable level of uptake of siHYB gene silencing molecules in a population of human cells *in vitro*. Furthermore, it appears that siHYB molecules are not susceptible to RNase H attack *in vitro*, even though RNase H has been shown to degrade RNA:DNA hybrids. This may help to explain the long-term gene silencing effects we have seen.

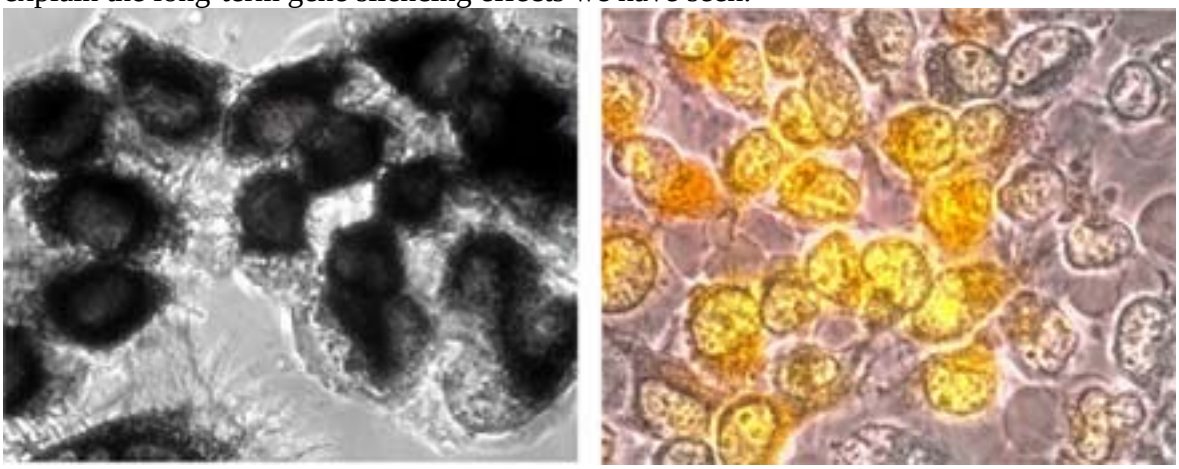

Figure legend: G6PD assay visualization. HeLa cells Cy3/Cy5 dual-labeled siHybrid against human glucose 6-phosphate dehydrogenase (G6PD) and assayed for G6PD activity. Purple precipitate is indicative of G6PD activity. Control cells (left) show strong G6PD activity. Transfected cells (right) show minimal purple precipitate (decreased G6PD activity) and Cy 3/5 co-localized within cells. Images taken on a Zeiss Axiovert 100 equipped with a CoolSnap HQ color camera and over-layed using Universal Imaging's Metamorph software.

Single-Cell Level Investigation of Cytoskeletal Response to External Stimuli

Amy Hiddessen [1,2], Allen Christian [2], David Clague [1,3], Thomas Huser [4], Alex Malkin [5], Luke Lee [6]

The cellular cytoskeleton plays a key role in fundamental cellular responses to external stimuli (i.e., chemical and physical changes in the extracellular environment) that contribute to both health and disease. Hence, an improved understanding of the cytoskeletal mechanisms which mediate cellular response to extracellular signals can lead to methods for preventing or treating illnesses ranging from the common cold to cancer. For example, little is known about single cell response to viral stimuli, particularly the host cell's cytoskeletal response to the initial interactions associated with viral docking and entry (Smith and Enquist 2002). An article published in *Science* just this year suggests that initial contact between human T cell leukemia virus-type 1 (HTLV-1) and an uninfected host cell induces rapid host cell cytoskeletal polarization as a necessary step for viral spreading, and may explain the relatively unknown mechanisms of cell-to-cell spread of other viruses, including HIV (Igakura et al. 2003). The objective of this research is to develop and apply a new experimental platform to characterize cellular cytoskeletal response to external stimuli of pathogenic origin. Using a combination of cytoskeletal labeling techniques and unique microfabricated cell environments (in collaboration with UC Berkeley and LLNL), single, living cells will be exposed to controlled concentrations of biochemicals derived from pathogens, and fluorescence microscopy will be employed to characterize the dynamic response of the cytoskeleton in real time. Ultimately, this research will further elucidate poorly understood cellular mechanisms of pathogenesis.

[1] Physical Biosciences Institute, [2] Biology and Biotechnology Research Program, [3] Engineering, [4]Chemical Biology and Nuclear Science Division & Medical Physics and Biophysics Division, [5] Chemistry and Materials Science, BioSecurity & Nanosciences Laboratory, Lawrence Livermore National Laboratory
[6] Department of Bioengineering, University of California at Berkeley

Single Molecule Studies of Chromatin

Chris Jeans[1], Michael Thelen[2], Alex Noy[3], Mike Colvin[1]

The DNA in eukaryotic cells is tightly packaged as chromatin through interactions with histone proteins to form nucleosomes. These nucleosomes are themselves packed together through interactions with linker histone and non-histone proteins. In order for processes such as DNA replication, DNA repair, and transcription to occur, the chromatin fiber must be remodeled such that the necessary enzymes can access the DNA. The structure of the chromatin fiber beyond the level of the single nucleosome and the structural changes which accompany the remodeling process are poorly understood. We are studying the structures and forces behind the remodeling process through the use of atomic force microscopy (AFM). This allows both high-resolution imaging of the chromatin, and manipulation of individual fibers. Pulling a single chromatin fiber apart using the AFM tip yields information on the forces which hold the structure together. Chromatin fibers have been isolated from chicken erythrocytes and Chinese hamster ovary cells. These have been analyzed by agarose gel electrophoresis, which shows the samples to contain a broad range of fiber sizes, from single nucleosomes to chains of 20 nucleosomes or more. SDS-PAGE shows the samples to contain only the core and linker histones, with no detectable contamination from cytoplasmic or non-histone nuclear proteins. AFM imaging has been carried out on chicken erythrocyte chromatin. In such images, single nucleosomes and short chains of nucleosomes are readily identifiable, and regions of linker DNA between nucleosomes can be seen. This work provides a strong basis for future experiments designed to stretch individual fibers with the AFM tip.

[1]Physical Biosciences Institute, [2]Biology and Biotechnology Research Program, [3]Department of Chemistry and Materials Science

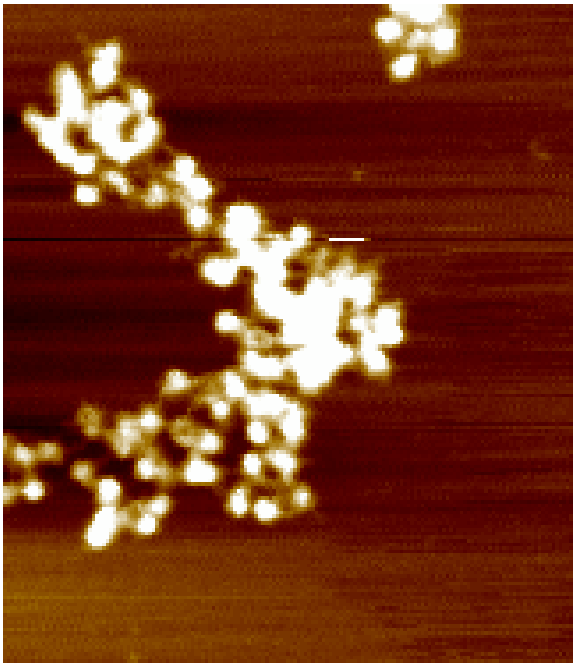

Figure legend: AFM image of a chicken erythrocyte chromatin fiber. Scale bar: 100 nm. Height scale: black to white = 0 to 15 nm. Individual nucleosomes are visible as white circles approximately 15 nm diameter. Linker DNA can be seen connecting the nucleosomes.

Application of SERS Nanoparticles for Intracellular pH Measurements

Ted Laurence[1], Chad Talley[2], Thomas Huser[2], Michael Colvin[1]

We are developing methods to measure concentrations in chemical microenvironments in cells and tissues using recently developed, functionalized metal nanoparticles (50-100 nm in diameter). Surface-Enhanced Raman Scattering (SERS) allows sensitive detection of changes in the state of chemical groups attached to single nanoparticles. A nanoscale pH meter has been tested in a cell-free medium, measuring the pH of the solution immediately surrounding the nanoparticles. We are applying single-particle detection methods to detect and characterize heterogeneity in the SERS signal. The nanoscale pH meters are being tested inside cells to determine optimal conditions and methods for these measurements. The initial focus of this project will be the measurement of intra- and extracellular pH in neuroblastoma cells lines that are known to have a large pH gradient across the cell membrane.

[1]*Physical Biosciences Institute, Biology and Biotechnology Research Program*
[2]*Chemistry and Materials Science Directorate*

Using femtosecond laser subcellular surgery as a tool to study cell biology

Nan Shen

Research on cellular function and regulation would be greatly advanced by new instrumentation using methods to alter cellular processes with spatial discrimination on the nanometer-scale. We present a novel technique for targeting submicrometer-sized organelles or other biologically important regions in living cells using femtosecond laser pulses. By tightly focusing these pulses beneath the cell membrane, we can vaporize cellular material inside the cell through nonlinear optical processes. This technique enables non-invasive manipulation of the physical structure of a cell with sub-micrometer resolution. We propose to study the role mitochondria play in cell proliferation and apoptosis, and ultimately the relation between the length of telomeres on a chromosome and cell senescence by selectively perturbing mitochondria or telomeres in living cells. Our technique will provide a unique tool for the study of cell biology.

Also contributing to the work: Mike Colvin and Thomas Huser.

Force Spectroscopy to Study Multivalent Binding in Protein-Antibody Interactions

Todd Sulchek

Understanding the complex interactions that occur between single molecules at surfaces of cells is important for many processes such as signaling and recognition. In particular, antibody-antigen interactions are important in biology and medicine because antibodies are our body's primary tool for fighting disease. Measuring the forces generated by these interactions should provide important information about mechanisms of antibody function. Here we propose to use the atomic force microscope (AFM) to measure the force of binding between the mucin 1 (Muc1) protein, which is indicative of certain cancers, and an antibody screened against Muc1. This binding interaction is the targeting mechanism in radioimmunotherapy -- a technique used to treat cancer tumors in which antiMuc1 is tethered to a radionucleotide. In this project, purified Muc1 protein is functionally attached to a surface using a spacer molecule. An antiMuc1 antibody is attached to an AFM tip. We have determined the unbinding force for one and more Muc1-antibody bonds. This information can help determine the optimal number of antibodies and tether length for use in radioimmunotherapy. Developing this force measurement strategy opens the door to studying other molecular interactions that occur at the interfaces of cells.

Significant Achievements

Much of the research within the Biology and Biotechnology Research Program crosses organizational boundaries. We have used general science topics as well as scientific groups as subdivisions to highlight the variety of research achievements accomplished during 2003.

Biodefense

- Developed new forensic DNA analysis techniques for the law enforcement community.

- Demonstrated hybrid interference technology for long-term gene silencing without the need for active transfection of the target cells in preclinical cancer tests.

- Designed and fabricated the Instrumented Cell integrated measurement platform, which enables single-cell analytical measurements.

- Identified potential biomarkers for presymptomatic detection of plague using a novel whole blood human host model, cell models, and animal models with 2-D DIGE multiplex proteomic characterization of both host and pathogen.

- Monitored expression of *Yersinia pestis* virulence factors and identified a possible new virulence induction mechanism using a real-time fluorescent based expression system.

- Silenced virulence factor genes in *Y. pestis* using an siRNA hybrid approach, and have begun functional and proteomic characterization of *Y. pestis* 'knockouts' to better characterize virulence and support future detection systems.

- Demonstrated identification and confirmation of virulent agents in aerosols using the Autonomous Pathogen Detection System (APDS).

- Demonstrated fully automated operation of the APDS in an airport and a subway station.

- Performed a human study to identify biochemical signatures for smallpox vaccination.

- Performed a SELDI-TOF-MS study of serum proteins in hemodialysis patients.

- Designed and built a computational system for determining unique protein targets for bacterial and viral pathogen detection, using a whole-proteome comparison approach.

- R&D 100 Award: For the development of the ***BASIS*** pathogen detection system.

- Designed and built a Monte Carlo system under intelligence agency funding to address the question "how much sequencing is enough for detection and forensic signature development." Results of this research have been submitted for publication and were used to guide 2004 national priorities for sequencing pathogens and near neighbors.

- Responded to requests from the Centers for Disease Control for assistance in determining smallpox, monkeypox, cowpox, and SARS signatures, using our unique computational capabilities, in response to ongoing world events.

- Completed and released to the public domain the entire sequence of *Francisella tularensis* LVS (live vaccine strain) an attenuated type B strain of the causative agent of tularemia.

- In collaboration with a European-American consortium based at the University of Uppsala, Sweden, the whole genome of *F. tularensis* Schu4, a fully virulent type A strain was also completed and released to the scientific community.

- Human dendritic cells obtained from different human donors have been infected with *Y. pestis* cells either possessing a functional type III secretory apparatus or lacking it. Molecular markers altered during the host in response to type III secretion function have been identified.

- Obtained NIH funding to study the role of unique chromosomal regions of *Y. pestis* in virulence.

- Developed a lambda red recombinase protocol for performing gene knockouts in *Y. pestis* and used it to knock out three *Y. pestis* unique chromosomal regions as candidate regions for new virulence genes.

- Characterized the global transcriptional response of *Y. pestis* to pH5.5 and 4.5 mimicking conditions found inside phagosomes. A gene group involved in this low pH response has been identified and the region has been deleted to ascertain its role in virulence.

- Completed eight and published to GenBank four sequences of the Filo and Arena RNA virus groups.

- Completed the genome sequence of *Brucella abortus* 2308.

- Patent applied for on a chip-based laser-mediated pathogen detection system.

- Photochemistry needed for the detector has been demonstrated using salmon sperm DNA as a model system.

- Proof of concept study under way to use accelerator mass spectrometry to find new peptide-based signatures in forensic samples of small size (in collaboration with Brian Souza and Steve Velsko).

- Collaborated with many federal, state and local partners to implement a national, environmental surveillance program that screens for the presence of bioterrorist agents (BIOWATCH).

Computational and Systems Biology

- Began work on the implementation of the Protein Model Database, a system to develop means for model quality assessment and structural comparative genomics. Protein Model Database will be a public facility extending the Protein Data Base (PDB), which addresses experimentally derived structures, to theoretical models.

- Accomplished a competitive extension of the "Center for Critical Assessment of Protein Structure Prediction" NIH funding.

- Following our previous NMR structural studies of DNA base excision repair proteins DNA Ligase III and XRCC1, we developed a new monoclonal antibody and used immunoprecipitation and affinity chromatography to determine the BRCT domain-specific interactions in XRCC1, PARP and DNA Ligase III.

- Towards a computational 3D model of cellular transport, we used laser scanning confocal microscopy to localize and quantitate the density of gap junctions, tight junctions and IP3 receptors of epithelial cells.

- As an ongoing project to characterize mutations that we identified in the Chinese hamster ovary XPB nucleotide excision repair protein, we examined the impact of these changes in protein sequence using molecular modeling, and in gene expression using *in situ* rolling circle amplification of mutant transcripts.

- Completed a pair of 2 ns molecular dynamics (MD) 50,000 atom simulations of a DNA sliding clamp protein in the presence and absence of DNA.

- Developed a physical theory to relate molecular simulations to experimentally measured decreases in thermal stability of mutant proteins found in the human population.

- We built a homology model of human P450 CYP1A2 and docked caffeine and PhiP into the active site. The homology models of CYP1A2 were found to be stable enough for molecular dynamics simulations and one of the models can accommodate caffeine in a stacked conformation over the heme (in agreement with NMR data). Initial docking calculations supported the rank order of PhiP mutation efficiency.

- Using molecular dynamics, we studied the solvation effects of DNA encapsulated in a single-walled carbon nanotube. Within a 30 A nanotube with the counterions excluded, the inserted DNA undergoes a severe bend similar to one found for DNA bound to the TATA-box binding protein.

- Using quantum calculations, we studied the chemical structures of DOTA chelating yttrium and DOTA chelating indium. Binding of a single water molecule to fill out the octahedral coordination around the metal ion does not affect the Y-DOTA structure but coordination state and geometry of In-DOTA changes.

Genome Biology

- Major participants in completion of the sequence and annotation of human chromosome 19 (soon to be published in the journal *Nature).*

- Discovered a new insulator element that is a likely imprinting control region (ICR) for chromosome 19 imprinted genes and established a role for transcription factor YY1 in its parent-of-origin specific function.

- Developed a new computational tool called *zPicture* to facilitate dynamic display and manipulation of conserved elements in genome alignments.

- Documented the mechanisms underlying lineage-specific expansion and divergence of zinc-finger transcription factor genes, analyzing for the first time the evolutionary forces operating on members of a specific gene cluster in detail.

- The I.M.A.G.E. Consortium continues to provide arrayed cDNA resources to the community through a network of distributors. We have arrayed over eight million cDNAs from seven species (832,000 cDNAs in 2003), resulting in over five million ESTs submitted to Genbank (dbEST) so far.

- As part of the Mammalian Gene Collection (MGC) project, we have re-arrayed nearly 88,000 (31,000 in 2003) predicted full-length genes and made them available to the research community. The full-length sequences have been submitted to Genbank. At this time 71% of human and 68% of mouse known genes have a full-length I.M.A.G.E. clone representative from this project.

DNA repair

- Accomplished detailed mapping of the domains for the protein interaction between XRCC2 and RAD51D. The findings provide insight for the structure-function studies on the DNA repair proteins.

- Discovered peptides that specifically bind RAD51 paralogs (RAD51B, C, D, XRCC2) by using phage display technique. These peptides have potential to be used in cancer therapy.

- Elucidated the differential roles of RAD51 paralogs (XRCC2) in S-phase RAD51 focus formation induced by DNA replication arrest. These findings help to establish a new concept that different pathways of homologous recombination are associated with stalled replication forks.

- Showed that a null mutation in the Fanconi anemia *FANCG* gene in CHO hamster cells causes an increased rates of gene amplification measured at the *DFHR* and *CAD* loci.

- Developed a molecular model that accounts for the chromosome instability and reduced *hprt* mutation rates seen in Fanconi anemia (FA) cell lines. The model invokes the Fanconi anemia proteins as acting to stabilize DNA replication forks that are blocked at crosslinks and other DNA lesions. In FA mutant cells, collapse of replication forks causes chromosomal breakage and leads to cancer-promoting rearrangements.

- Introduced the human Fanconi Anemia *FANCG* gene into an immortalized *fancg* mutant fibroblast line and showed that the gene produced full complementation. The complemented and parental cells provide a genetically matched pair that will be used in a wide variety of studies to identify the function of FANCG protein.

- Used DNA repair-deficient mutant lines of CHO cells to show the contributions of the double-strand break repair pathways to radiation sensitivity through the cell cycle. Homologous recombinational repair acts primarily during S phase, whereas nonhomologous end joining acts throughout the cell cycle.

- Showed that the human XRCC3 and Rad51C Rad51 paralogs, which act in homologous recombinational repair, form a dimer that separates when ATP binding occurs. This finding suggests that ATP may exert a regulatory role on these proteins during recombinational repair. Mutations in XRCC3 that prevent ATP binding or hydrolysis cause complete loss of repair activity.

- Showed that XRCC3 and Rad51C can be overexpressed as a stable dimer complex in bacteria in quantities that may be sufficient for purifying the complex for protein crystallization and structural studies.

- Determined that RAD51B, RAD51C, RAD51D and XRCC2 form a stable, protein complex *in vivo*, devoid of RAD51.

- Used yeast two-hybrid analysis and molecular modeling to map the interaction of the RAD51B-RAD51C interaction.

- Generated stable cell lines that over-express wild-type and mutant RAD51B and RAD51C proteins.

- Developed the technology to produce recombinant human proteins in a miniaturized, baculovirus-based system.

- Published a book on novel protein array and proteomics technologies and their applications.

- Developed miniaturized immunoprecipitation techniques in baculovirus in a 96-well format.

- In collaboration with the IMAGE Consortium, developed an automated, high-throughput, open reading frame (ORF) library construction process for ORFs from genomic DNA of the microbes *Sinorhizobium meliloti* and *Ralstonia metallidurans*.

- Performed comparative immunoassays to determine limits of detection, reproducibility, and potential for miniaturization between the Luminex™ system and protein microarrays.

- Identified the binding domains of the Rad51B-Rad51C-Rad51D-Xrcc2 and Rad51C-Xrcc3 complexes to further characterize the interaction of these proteins and propose a model for the three-dimensional architecture of these paralog complexes.

- Demonstrated that the DNA repair protein RAD51B, and not its cognate partner RAD51C, interacts with histones and not nucleosomes.

- Identified novel protein interactions between RAD51B and core histones, providing new understanding of the cellular role of RAD51 paralogs.

Mutagenesis and carcinogenesis

- Established that the dietary carcinogen PhIP causes lobe-specific histologic changes in the prostate of rats after 20 weeks of exposure in a study to develop a preclinical animal model for prostate cancer. These changes may provide clues as to the mechanisms involved in prostate cancer development and help establish molecular targets for prostate cancer prevention.

- Showed that the level of PhIP-DNA adducts formed in human colon may be related to age and polymorphisms in several enzymes involved in PhIP metabolism.

- Showed that accelerator mass spectrometry (AMS) can detect drug levels in cancer cells with a million-fold increase in sensitivity over traditional techniques which provides supporting evidence for developing an AMS-based method for quantifying drug levels in target.

- Showed that the anti-estrogen tamoxifen causes DNA adducts in the endometrium of women and may therefore be involved in the initiation of cancer.

- Structurally identified a protein adduct formed by PhIP which may provide a useful biomarker of exposure.

- Differentiated the mass spectra of breast cancer cell homogenates acquired using Time-of-Flight Secondary Ion Mass Spectrometry. The spectra were analyzed and grouped using multivariate statistical analysis. This technique allowed us to discriminate different cell lines, different regions within a cell and cells treated with a proliferative agent. Using these results, we can now develop methods 1) for early detection of cancerous changes in cells and 2) to identify and classify cancer cells according to their primary tumor or malignant potential.

- Determined that the metabolism of the food carcinogen, PhIP, varies in individuals over time in a study of 12 volunteers. This finding suggests that lifestyle factors may influence carcinogen metabolism and individual cancer risk.

- Completed a systematic study of the cooking parameters that affect heterocyclic amine formation in cooked meat. This comprehensive, statistician-designed study will allow us to propose cooking strategies aimed at lowering exposure to cooked meat carcinogens.

- Determined that dietary fiber affects the bioavailability of heterocyclic amines. Using our *in vitro* digestion model, we showed that fiber absorbs heterocyclic amines released during digestion of cooked meat. This study may also lead to strategies for lowering heterocyclic amine exposure in humans.

- In our continuing analysis of melanoma in LLNL employees, we have determined that the melanoma surveillance effort begun in 1984 is associated with zero melanoma mortality in the period from 1984 to 1997. This is a significant reduction below expected mortality based on California mortality rates.

Genetic Susceptibility

- Developed histone protein microarray assays for characterizing structural and functional interactions for the SMARCAL1 protein, which is involved in the childhood disease Schimke immuno-osseous dysplasia. These results identify the importance of nucleosomal interactions for SMARCAL1 function in organization of chromatin.

- Completed a comparison study of multiplexed techniques using liquid arrays for detection of proteins, establishing several guidelines for selecting between protein microarrays and liquid array technologies for diagnostic uses.

Reproductive and Developmental Biology

- Showed that exposure of male mice to the chemotherapeutic agent etoposide has major chromosomal effects on their meiotic cells; exposure to etoposide may also have long lasting effects on the frequencies of sperm with chromosomal structural aberrations.

- Showed that varying the hormonal treatment of female mice can affect fertility by creating an hostile uterine environment that is detrimental to sperm function and motility.

- Showed that DNA single-strand breaks and alkali labile sites, but not DNA double strand breaks, are significantly increased in sperm of older men.

- Microarray analysis of gene expression in mouse oocytes identified ~60 differentially expressed genes in the T-stock strain, which is known for its reduced ability to repair sperm DNA damage.

- Showed that paternal exposure to various mutagens alters the gene expression pattern in the early embryo in a mutagen-specific manner.

Radionuclide and Radiation Biology

- Showed that low dose ionizing radiation (10cGy or less) modulated the transcription of a set of genes that differed substantially from those induced after high-dose exposures in human lymphoblastoid cell lines derived from multiple individuals. These findings suggest that the low-dose response induced cellular response pathways that were qualitatively different from those employed at higher doses.

- Conducted a genome-scale survey of the effects of low dose ionizing radiation on gene expression changes in the brain of irradiated mice. Low-dose exposure resulted in early changes in genes involved in signal transduction, ion regulation and synaptic signaling. Late changes involved a number of metabolic pathways including myelin and protein synthesis. There were also changes in expression of genes involved in stress response, cell-cycle control, and DNA synthesis/repair. The brain appears to respond to low dose irradiation by evoking protective as well as reparative mechanism while down-regulating genes involved in neural signaling.

- Identified and characterized specific genes whose expression patterns were associated with cytogenetic radio-adaptive response, which occurs when a low-dose of ionizing radiation reduces the cytogenetic damage from a subsequent higher dose of irradiation. We found that the radioadaptive responses may be attributed to increased expression of DNA damage sensing genes, while genes that drive cellular proliferation were down-regulated.

- Completed a pilot proteomic-scale survey on the effects of ionizing radiation on human lymphoblastoid cells, and found that the highest differential levels involved proteins that have been associated with heat shock-related stress responses.

- Established a method for preparing samples for analysis of tritium content by accelerator mass spectrometry (AMS). This method was used to analyze tree rings and hence establish environmental tritium exposures.

- Improved the web-based technology for querying of multiple bioinformatics sources to investigate the roles of co-regulation of genes in various types of radiation responses. Our web-based technology provides the biologist with a quickly adaptable way to identify genes and functional pathways that are coincidentally related to gene expression patterns.

- Established a Lab Information Management System (LIMS) for the Microarray Collaboration Group to hold microarray gene expression data sets and related experimental data for access to members. This website provides user-friendly access to several microarray laboratory protocols, free bioinformatic software, microarray service and published data sets.

Molecular Biophysics

- The first set of synthetic, high affinity ligands (SHALs) were developed that bind the unique region of the human lymphoma tumor cell receptor HLA-DR10. Experiments with cultured cells and tumor tissue sections have shown that one of these SHALs binds to large and small cell lymphoma cells with nM affinity but not to normal cells.

- In an ongoing effort to understand how adduct conformation affects cellular responses to DNA damage, we have determined by NMR a novel structure of a DNA adduct (O^6-[4-oxo-4-(3-pyridyl)butyl]guanine in an 11mer duplex) derived from a tobacco-specific carcinogen.

- Disease associated variants of human Ape1 were found to be thermally destabilizing by scanning calorimetry and enzyme assays. Expression clones and variant constructs of three other DNA repair proteins were constructed for parallel analyses.

- New methods were developed for the high-throughput production of proteins by over-expression. The primary focus included vector development, high-throughput protein expression screening and characterization by size-exclusion chromatography/on-line light scattering.

- As part of the NIH TB Consortium project, we have determined the crystal structure of the *Mycobacterium tuberculosis* rmlC epimerase (Rv3465), a promising drug target structure in the rhamnose pathway.

- Using a combination of optical trapping and Raman spectroscopy, we have developed a method that makes it possible to detect differences (or changes) in the protein and nucleic acid content of individual living cells in real time.

- We have developed Bayesian probability estimates for the resolution dependence of crystallographic unit cell contents (Matthews probabilities), and the protein isolectric point as a predictor for protein crystallization probability, which has also for the first time deconvoluted the empirically observed protein crystallization pH distribution and provided a rational biophysical basis for pH preference. We have also introduced propensity analysis and applied Bayesian principles to machine learning in the field of protein crystallization analysis.

- Developed and implemented at Texas A&M University an effective electron density map improvement and structure validation on a Linux multi-CPU web cluster: The TB Structural Genomics Consortium Bias Removal Web Service.

Bioinformatics

- Support of the Joint Genome Institute production sequencing facility's ongoing operations and ramp to nearly 2 billion basepairs of sequence generated per month.

- Contributions to whole-genome analysis and annotation jamborees and numerous journal publications.

- Involvement in the sequencing, analysis, and annotation of human chromosomes 5, 16,and 19.

- Continuous support and improvements to the I.M.A.G.E. Consortium, with a recent release of IMAGEne version 4.4.

- Participation with Microarray Collaboration Group in the design and analysis of microarray experiments.

- Contributions to various BBRP proposals and grant applications.

- Support of analysis and annotation of environmental restoration-related microbial genome.

- Design and development of adminstrative databases and interfaces.

- Support of comparative analysis and annotation of syntenic chicken/human sequences.

2.0 Program Organization, Facilities and Resources

Organizational Structure

The Program is organized with seven functional units as shown in Figure 1. These units are designed to match the new opportunities in homeland security, computational biology, instrumentation, and particularly to take advantage of the new information and tools that are becoming available as we move from a focus on acquiring DNA sequence data to understanding genetic and biochemical function, diversity, and fundamental biological strategies. The program units include the:

Genome Biology Division
Computational and Systems Biology Division
Biodefense Division
Health Effect Genetics Division
Molecular Biophysics Group
Environmental Microbiology Group
Bioinformatics Group

Environmental Microbiology is a new thrust area closely tied to the interests of the DOE Office of Science Genomes to Life program. With the support of the Laboratory, we are actively recruiting strong researchers and scientific leaders to develop this expertise in support of the DOE/OBER objectives. The organizational structure both depends on and builds on BBRP's tradition of 'invisible' organizational boundaries; e.g. scientists must work together across these boundaries in a highly interactive fashion with colleagues within the program and across the Laboratory broadly.

Within the BBRP Program Office, Alan Casamajor acts as Assistant to the Associate Director (AAD) for Facilities. Patsy Gilbert serves as Assurance Manager to ensure the highest levels of safety and compliance with regulatory requirements. The AAD for Resource Management, Cynthia Gardner, is responsible for our business operations and finances. Peggy Biggs serves as the Program Administrator and manages the administrative, human resources, and computer desktop support functions for the Program.

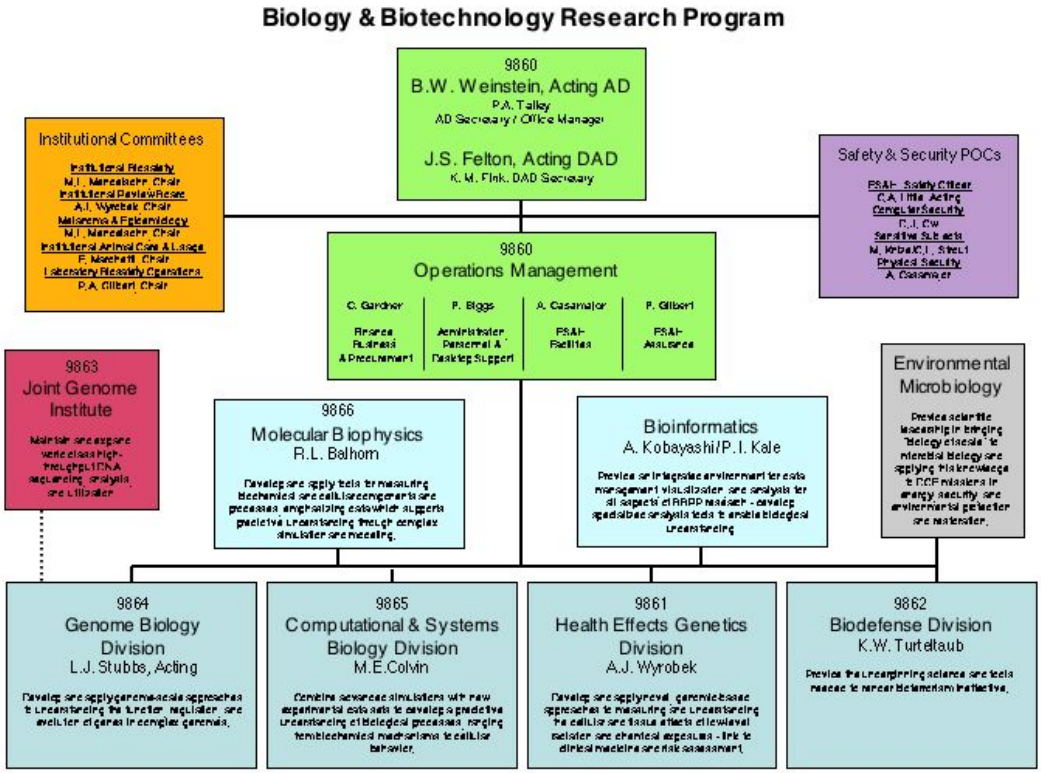

Figure 1A. Organizational chart for the period covered by the report, 2003.

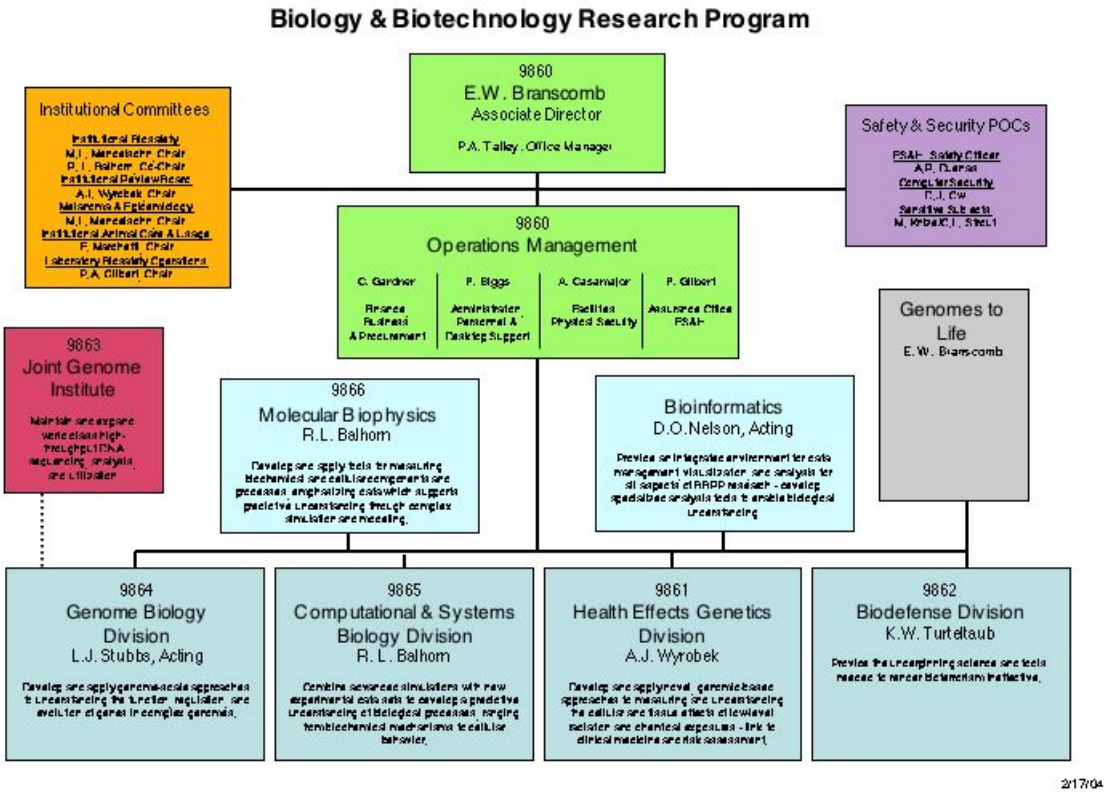

Figure 1B. Current organization structure, 2004.

Program Reviews and Advisory Committee

A variety of methods are used to assess the quality of our science and technology. At the project level, each federally funded effort is subject to a peer review process specific to its sponsor. For NIH and DOE, peer review is required for initial funding and for renewal. Every project is externally reviewed at least once every three years. At the program level, we undergo a bi-annual internal review by the LLNL Director, several of the Associate Directors, other LLNL staff, and laboratory consultants. This review focuses on strategic planning, future scientific directions, intra-laboratory interactions, and workforce issues. BBRP's Laboratory Directed Research is reviewed both in BBRP and by an LLNL-wide committee. Our BBRP Advisory Committee meets yearly to review the entire program on a three-year cycle. The standing membership of the Advisory Committee is shown in Table 1. Ad-hoc reviewers are included each year as experts in the specific areas being reviewed.

Table 1. Program Review Committee

Biology and Biotechnology Research Program Advisory Committee, 2003

Charles R. Cantor	Sequenom, Inc., *Chair*
Gilbert Chu	Stanford University Medical School
Fred F. Kadlubar	National Center for Toxicological Research
Hsing-Jien Kung	University of California, Davis Medical Center
Michael Levitt	Stanford University
Alex MacKerell	University of Maryland at Baltimore
Stephen A. Morse	Centers for Disease Control and Prevention
Frederick A. Murphy	University of California, Davis
Richard M. Myers	Stanford University School of Medicine
Kenneth H. Nealson	Jet Propulsion Laboratory
David A. Relman	Stanford University / Palo Alto Health Care
John A. Tainer	The Scripps Research Institute
Snorri Thorgeirsson	National Cancer Institute
James M. Tiedje	Michigan State University
Barbara J. Wold	California Institute of Technology
Peter Jahrling (ad hoc)	United States Army Medical Research Institute for Infectious Disease
Thomas Blumenthal (UC S&T Panel Representative)	University of Colorado School of Medicine

Lastly, the University of California, our contractor, conducts annual program reviews. The Science and Technology Council composed primarily of university faculty and appointed by the University President reviews the quality of science conducted at UC's three national laboratories. This Council meets with laboratory staff at least annually and receives written input from the Program Advisory Committee.

Budget Information

We draw upon several funding sources for our Program. While the Department of Energy remains our primary sponsor, in the past several years we have attracted a diverse assortment of funding sources. Funding amounts for each category are detailed in Table 2. The percentage funding by major funding source is shown in Figure 2. Appendix V details funding from DOE as well as non-DOE sources and from the LLNL Lab Directed Research and Development Program.

Table 2. Current funding sources

Sponsor	K$	Funding %
DOE, Office of Science/Office of Biological & Environmental Research	11,241	28.1
DOE, Office of Non-Proliferation and National Security	11,449	28.7
National Institutes of Health (four institutes)	8,621	21.6
Los Alamos National Laboratory	632	1.6
LLNL Laboratory Directed Research and Development	5,954	14.9
LLNL General and Administrative	187	0.5
Army	845	2.1
Federal, Other	255	0.6
University of California, Berkeley	407	1.0
University of California, Breast Cancer Research Program	162	0.4
University of California, UC Discovery Grants	83	0.2
Industry	125	0.3
TOTAL	**39,962**	**100.0**

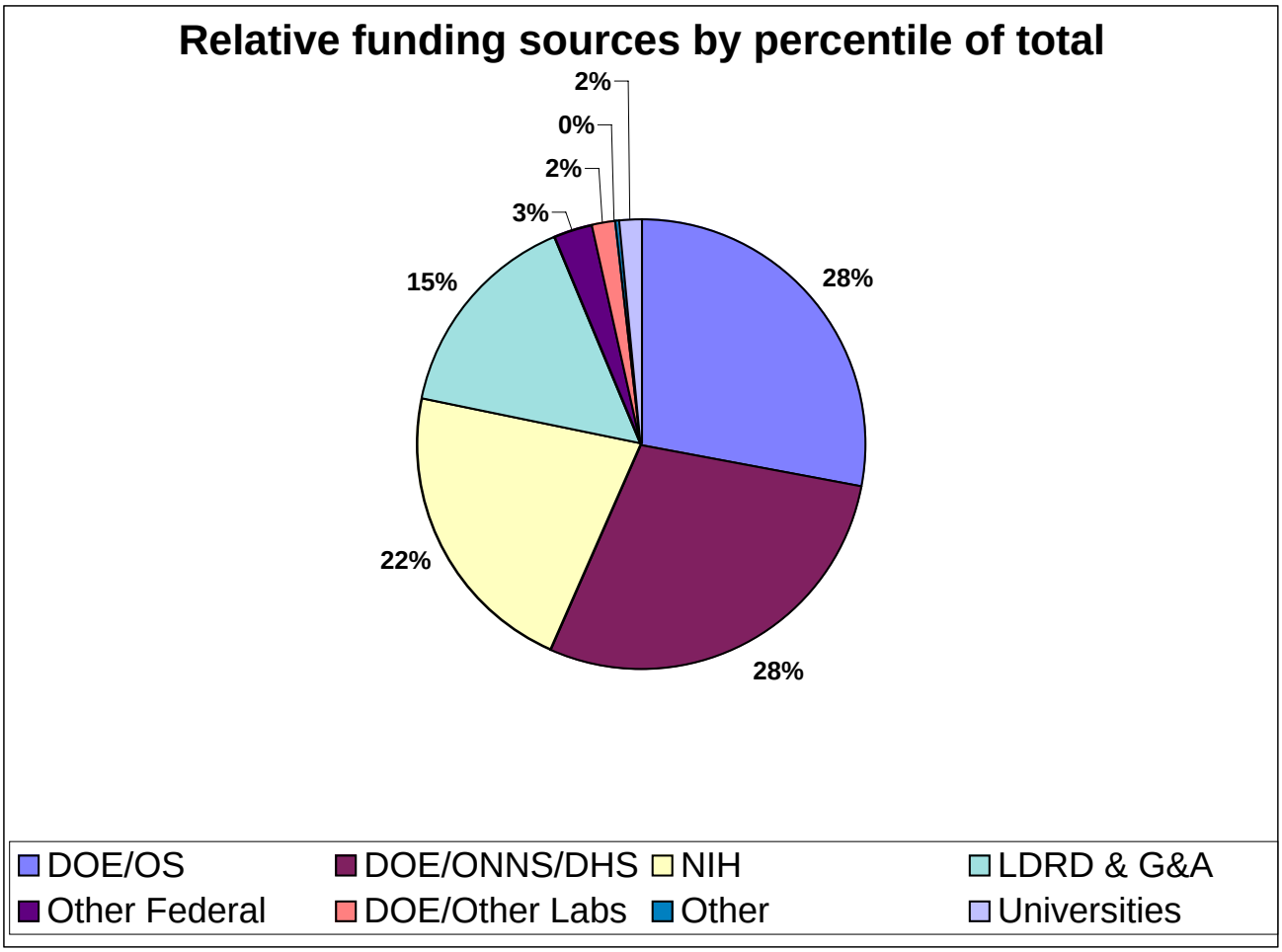

Figure 2. Relative funding sources by percentile of total

Personnel Update

We draw from a diverse talent pool and facilitate staff advancement through formal training and education programs, self-training, and mentoring. Table 3 presents a snapshot of the BBRP personnel by category. Additional information including a roster of principal investigators and their extramural activities is available in Appendices I and II.

Table 3. Staff by Category – October 2003

Personnel	FY 2002	FY 2003	FY 2004 (est)
Principal Investigators	30	28	32
Other Scientific Staff	91	110	110
Technicians	36	35	35
Post- Doctorates	17	26	30
Graduate Students	5	5	5
Undergraduates	32	30	30
Total Scientific Personnel	**211**	**234**	**242**
Professional Admin.	10	14	14
Clerical Admin. Support	19	17	18
Total Admin. Personnel	**29**	**31**	**32**
TOTAL	**240**	**265**	**274**

The Biology and Biotechnology Research Program provides educational and professional training opportunities for their staff, as well as to a wide array of students and faculty. The research and work experiences are designed to enhance the quality of science education at all levels and encourages diversity among people pursuing careers in the biomedical sciences. The programs integrate students and faculty into ongoing research projects and provide opportunities for them to expand their scientific and technical knowledge. The programs encourage students to continue their education in science. The summer programs allow students and faculty to extend and refresh their training and knowledge of areas of current research, and hence increase their ability to encourage students to pursue careers in science.

Year-round Programs

Science and Education Research Semester (SERS). This program, jointly funded by DOE and LLNL, provides undergraduates an opportunity to conduct an extended research project as a member of an ongoing research program for an academic semester.

Student Technology Experience Program (STEP). This program provides meaningful work experience to high school and college students, to encourage them to pursue further education in science, engineering or business.

Biology & Biotechnology Research Student Technical Program. This program is designed to provide opportunities for undergraduate college students to pursue a specific research project relevant to their studies.

Student Participating Guests. These students use lab facilities and/or participate in lab work as part of their academic studies for high school, an undergraduate, or graduate degree.

Participating Teacher/Guests. These teachers use lab facilities and/or participate in lab work to keep abreast of current research and improve classroom instruction.

Summer Programs

Summer Research Internship Program. This program matches teachers with research projects where they learn about new areas of science and technology and identify materials that can be incorporated in their curricula.

LLNL Summer Employment Program. This program coordinates the summer employment opportunities for students throughout the Laboratory.

Associated Western Universities Summer Program (AWU). This cooperative research program promotes and coordinates research and training in energy-related science and engineering for faculty and students of the Associated Western Universities.

Facilities

The major building housing the Biology and Biotechnology Research program is a 68,000 gross square foot, one-story structure containing 55 laboratories (24,900 sq. ft.), 83 offices (12,400 sq. ft.), auditorium, stock room, and administrative offices. Other laboratories (17,500 sq. ft.) and offices (11,000 sq. ft.) are located in fifteen nearby buildings and trailers, including special facilities for housing small animals, laboratories for handling virulent pathogens, a radiation exposure facility (a sealed 4,260 Ci ^{137}Cs source), an analytical chemistry laboratory, x-ray diffraction systems, and a high-level carcinogen laboratory (see Table 4).

BBRP is projected to grow at a rate of 5% per year for the next several years and to develop new areas of research involving work with pathogenic microbes. We have begun construction on a new, 1600 sq. ft. Biosafety Level 3 laboratory facility. This new facility would provide workspace for two funded activities that are currently using BSL3 facilities at remote locations, and would improve access to additional research dollars in the bio-defense area. We continue exploring options of adaptive reuse of existing buildings at LLNL and new construction to provide for near-term program expansion.

Table 4. BBRP Facilities

Facility	Use	Office (Sq.ft.)	Laboratory (Sq.ft.)
Bldg. 361	laboratory & office	12,382	24,863
Bldg. 362	laboratory & office	633	1,937
Bldg. 363	laboratory	0	1,119
Bldg. 364	animal care & laboratory	0	7,057
Bldg. 365	laboratory & office	188	2,969
Bldg. 366	laboratory & office	344	1,601
Bldg. 367	office	305	0
Bldg. 373	storage	0	0
Bldg. 376	shop & office	90	0
Bldg. 377	laboratory & office	349	2,823
Trlr 3629	office	1,482	0
Trlr 3649	library & computations	0	0
Trlr 3703	office	6,469	0
Ttlr 3751	office	1,708	0
Trlr 3775	office	1,071	0
Trlr 3777	office	4,610	0
	Totals:	**29,828**	**42,369**

Compliance with Federal Regulations

Lawrence Livermore National Laboratory has a long standing tradition of concern for the welfare of animals used in research, in the protection and confidentiality of human subjects, and in the careful use of biohazardous materials and biological organisms. Permanent institutional committees supported by the laboratory exist in each area. The Laboratory Director appoints the chair of each committee.

Since 1973, the Animal Care Facility has been in full compliance with the Animal Welfare Act, the U.S. Public Health Service, the National Institutes of Health, the Office of Laboratory Animal Welfare, and the University of California regulations governing the use of animals. In 1976 we established an active Institutional Animal Care and Use Committee (IACUC), currently chaired by Francesco Marchetti, whose membership includes representatives from our scientific staff and the local community. Our Animal Welfare Assurance, which dictates the institutional compliance requirements for LLNL, was recently renewed and is effective until July 31, 2007. We are also fully accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care International. Our accreditation has been recently renewed and is in effect until 2005.

The Institutional Review Board (IRB) was formed at LLNL in 1974 to provide institutional assurance that all research activities involving human participants were being conducted under internationally recognized ethical principles, in compliance with U.S. federal regulations, and in adherence to the policies of the U.S. Department of Energy, Department of Health and Human Services (DHHS), State of California, and University of California. The LLNL IRB has the responsibility to identify and review all human research activities conducted by or involving LLNL employees to assure that the research is justifiable, and that all human participants are protected from unnecessary harms and risks. The LLNL IRB operates under general provisions of the Code of Federal Regulations, 10 CFR Part 745, revised June 1991, and holds a Federalwide Assurance (FWA #00004274) with DHHS, Office for Human Research Protections. The FWA dictates the specific institutional requirements for maintaining federal compliance and remains in effect until February 26, 2006. The LLNL Director provides the administrative support for the IRB office (web address: http://www.llnl.gov/HumanSubjects) and the review of the individual research protocols is conducted by the IRB which consists of volunteer scientific and non-scientific experts, including institutional and Livermore community members. Andrew J. Wyrobek is currently the manager of the IRB office and chair of the IRB (wyrobek1@llnl.gov).

The Institutional Biosafety Committee (IBC) was established at Lawrence Livermore National Laboratory in 1991 to ensure compliance with regulations concerning work with recombinant DNA or human, animal, and plant pathogens. These regulations include the NIH Guidelines for Research Involving Recombinant DNA Molecules, April 2002, and the CDC-NIH, Biosafety in Microbiological and Biomedical Laboratories, 1999 (Revision 4). All work in the cited subject areas must be approved by the IBC prior to initiation. In the past year, a highly cross-referenced web site (http://ibc.llnl.gov) was

created to help staff to understand and meet the Committee's requirements. The Committee consists of scientists, physicians, lawyers, and community members, and is chaired by Mort Mendelsohn and Rod Balhorn.

3.0 Appendices

Appendix I: Biology and Biotechnology Research Program PhD Staff and Project Managers

Principal Investigators / Project Managers

Agron, Peter	PhD	University of California, San Diego	Biology
Albala, Joanna S.	PhD	Yeshiva University, Bronx, NY	Pathology
Ascher, Michael S.	MD	Harvard University, Boston, MA	Medicine
Balhorn, Rodney L.	PhD	University of Iowa, Iowa City	Biochemistry/Chemistry
Barsky, Daniel	PhD	University of Illinois, Urbana	Biophysics
Beckwith, William	PhD	Wesleyan University, Middletown, CT	Molecular Biology
Beernink, Peter	PhD	Boston University, Boston, MA	Biology
Chain, Patrick	MS	McMaster Univ., Ontario, Canada	Molecular Biology
Christian, Allen	PhD	Colorado State University, Ft. Collins	Chemical Engineering
Coleman, Matt	PhD	Boston University, Boston, MA	Molecular Biology
Cosman, Monique	PhD	New York University, New York	Physical Chemistry
Detter, John	PhD	University of Florida, Gainesville	Molecular Biology
Dingley, Karen	PhD	York University, England	Biochemistry/ Environmental Toxicology
Felton, James S.	PhD	State Univ. of New York, Buffalo	Molecular Biology; Zoology
Fidelis, Krzysztof A.	PhD	University of Oklahoma, Tulsa	Physical Chemistry
Garcia, Emilio	PhD	University of California, Davis	Microbiology
Gardner, Shea	PhD	University of California, Davis	Ecology
Gordon, Laurie	BS	Graceland College, Lamoni, IA	Biology
Henderson, Paul	PhD	Georgia Inst. of Technology, Atlanta	Chemistry
Jones, Irene	PhD	University of Illinois, Urbana	Cell Biology
Kale, Patricia *	BS	University of California, Berkeley	Electrical Engineering/ Computer Science
Kim, Joomyeong	PhD	Louisiana State Univ., Baton Rouge	Molecular Biology
Knize, Mark	BA	Calif. State Univ. Stanislaus, Turlock	Biology
Kobayashi, Art *	MS	Calif. State Univ. Sacramento	Computer Science
Krishnan, Viswanathan	PhD	Indian Inst. of Science, Bangalore	Physics
Kryshtafovych, Andriy	PhD		
Kulp, Kris	PhD	University of California, Davis	Pharmacology
Langlois, Richard G.	PhD	Columbia Univ, New York; UCBerkeley	Chemistry
Lightstone, Felice	PhD	Univ. of California, Santa Barbara	Chemistry
Liu, Nan	PhD	St. Andrews University, UK	Medicine
Loots, Gabriela	PhD	University of California, Berkeley	Microbiology
Lowe, Xiu	MS	California State Univ., Hayward	Molecular Biology
Lu, Xiaochen	MA		Pathology
Lucas, Susan	BA	Oregon State University, Corvallis	Biology
Manohar, Chitra	PhD	Indian Inst. of Science, Bangalore	Molecular Biology
Marchetti, Francesco	PhD	University of Rome, Italy	Biological Sciences
McCutchen-Maloney, S.	PhD	Texas A&M Univ., College Station	Chemistry
Mendelsohn, Mortimer L.	MD/ PhD	Harvard Medical School, Boston MA	Biophysics
Messenger, Sharon	PhD	University of Texas, Austin	Zoology
Moore, Dan	PhD	University of California, Berkeley	Statistics

Nelson, David O. *	PhD	University of California, Berkeley	Statistics
Ovcharenko, Ivan *	PhD	Institute of Chemical Kinetics and Combustion, Novosibirsk, Russia	Physics/Mathematics
Ow, David *	MS	Calif. State Univ. Stanislaus, Turlock	Computer Science
Prange, Christa	BS	St. Mary's College, Moraga, CA	Biology
Radnedge, Lyndsay	PhD	University of London, UK	Bacteriology
Rupp, Bernhard	PhD	University of Vienna, Austria	Chemistry
Segelke, Brent	PhD	University of California, San Diego	Chemistry
Skowronski, Evan	PhD	Loma Linda Univ., Loma Linda, CA	Microbiology
Slezak, Tom *	MS	University of California, Davis	Computer Science/ Mathematics
Smith, Kimothy	DVM/ PhD	Louisiana State Univ, A&M College, Baton Rouge	Biology
Stubbs, Lisa J.	PhD	University of California, San Diego	Molecular Biology
Tebbs, Robert	PhD	Wayne State Univ., Detroit, MI	Chemistry
Thelen, Michael P.	PhD	Cambridge University, England	Biochemistry
Thompson, Lawrence	PhD	Univ. Texas Health Center, Houston	Biophysics
Turteltaub, Kenneth	PhD	Iowa State University, Ames	Toxicology
Venclovas, Ceslovas	PhD	M.V. Lomonosov Moscow State Univ., Russia	Chemistry
Vitalis, Elizabeth	PhD	University of California, San Francisco	Biomedicine
Weinstein, Berthold W.	PhD	University of Illinois, Urbana	Physics
Wilson, Wendy	PhD	Cornell University, Ithaca, NY	Plant Pathology
Wyrobek, Andrew J.	PhD	University of Toronto, Canada	Medical Biophysics
Zhou, Carol *	PhD	University of Missouri, Columbia	Biological Sciences

** home organization is not BBRP*

Appendix I: Biology and Biotechnology Research Program Senior Staff, cont'd

Post-Doctoral Staff

Alegria-Hartman, Michelle	University of California, Davis	Genetics
Amer, Halima	Imperial College of Science and Technology, London	Chemistry
Bennion, Brian	University of Washington, Seattle	Biomedicine
Chromy, Brett	Northwestern University, Chicago, IL	Chemistry
Collette, Nicole	University of California, Davis	Genetics
Dover, Nir	Hebrew University, Jerusalem, Israel	Physiology/Biotechnology
Dugan, Lawrence	Colorado State Univ, Ft. Collins, CO	Radiobiology
Elso, Colleen	University of Melbourne, Australia	Genetics
Forde, Cameron	University of Toronto, Canada	Chemistry
Hah, Sang Soo	Seoul National University, Korea	Organic Chemistry
Hamilton, Aaron	University of California, Riverside	Biology
Hiddeson, Amy	Univ. of Pennsylvania, Philadelphia	Chemical Engineering
Hinz, John	University of Utah, Salt Lake City	Biology/Oncology
Huntley, Stuart	Washington State Univ, Pullman	Microbiology
Jeans, Christopher	Imperial College of Science, London	Biochemistry
Kim, Ken*	University of California, Berkeley	Physics
Lakshmanan, Nagarajan	Concordia University, Montreal, Canada	Molecular Biology
Lau, Edmond	University of California, Santa Barbara	Chemistry
Laurence, Ted	University of California, Berkeley	Physics
Pesavento, Joseph	Baylor Coll. of Medicine, Houston, TX	Biochemistry
Redding, Kellie	Ohio State University, Columbus	Molecular Biology
Sawika, Dorota	Boston University, Boston, MA	Computational Chemistry
Shen, Nan	Harvard University, Boston, MA	Experimental Physics
Sulchek, Todd	Stanford University, Stanford, CA	Experimental Physics
Yamada, Nazumi	Univ. of N. Carolina, Chapel Hill, NC	Pathology
Yoshikawa, Daniel	University of Rochester, Rochester, NY	Chemistry/Pharmacology
Zhang, Celia	Massay University, New Zealand	Biology/Genetics

**home organization is not BBRP*

Appendix II: Staff Extramural Activities

Major Awards and Honors

J. Albala	University of California Davis, Biotechnology Award 2003
P. Beernink	University of California Davis Cancer Research Symposium: Best Poster Award
K. Kulp	IDEA Award, California Breast Cancer Research Program
	National Institutes of Health/National Center for Complimentary Medicine Award
R. Langlois	BBRP Directorate Award for success of the Chernobyl P01 Project, January 2003
E. Skowronski	R&D 100 Award (BASIS)
	NAI Directorate Award (Exotic Newcastle Disease monitoring)
	NAI Directorate Award (BIOWATCH)
B. Weinstein	Honored Alumnus, Brigham Young University College of Physical and Mathematical Sciences
W. Wilson	NAI Directorate Award (BIOWATCH)
N. Yamada	Young Investigator Award, Gordon Research Conference on Radiation Oncology, January 2003
	Radiation Research Society Scholar-in-Training Award, 12 International Congress of Radiation Research, August 2003

Invited lectures

J. Albala	PEPTALK: Protein Expression, San Diego, CA, January 2003
	Genomes to Life Workshop - Protein Production (Facility I), Argonne, IL, May 2003
	Genomes to Life Workshop - Protein Characterization (Facility III), Atlanta, GA, June 2003

J. Albala, cont'd	Western Eye Research Conference, Laguna Beach, CA, September 2003
	3rd Annual Ontario Microarray Network Technology Symposium, Toronto, Canada, November 2003
P. Beernink	Trends in Biocalorimetry Symposium, Burlingame, CA
A. Christian	International Society of Analytical Cytometry
	Biodefense to Cancer, University of California Davis Cancer Center, CA
	University of California Los Angeles/University of California Riverside Toxicology meeting
B. Chromy	Human Proteome Organization (HUPO) 2nd Annual & International Union of Biochemistry and Molecular Biology (IUBMB) XIX Joint World Congress, Montreal, Canada, October 2003
M. Cosman	American Association for Cancer Research, Washington D.C, July 2003
	University of California Davis Cancer Center Ninth Annual Cancer Research Symposium, October 2003
	University of California Davis Cancer Joint Cancer Therapeutics-Biomedical Technology meeting, December 2003
J. Felton	American Association for Cancer Research (AACR) Molecular Epidemiology Conference, Kona, HI, January 2003
	Chair, Agricultural Health Conference and Symposium, Bethesda, MD, February 2003
	Environmental Mutagen Society (EMS) Colon Cancer Symposium, Miami, FL, May 2003
	International Environmental Mutagen Meeting, Florinopolis, Brazil, May 2003
	Reciprocal Meat Conference, University of Missouri, June 2003

J. Felton, cont'd	5th Congress of Toxicology in Developing Countries, Guilin, China, November 2003
	National Cancer Institute Special Workshop on Cancer Etiology, Washington, D.C., December 2003
K. Fidelis	Pacific Symposium on Biocomputing, Lihue, HI, January 2003
	European Conference on Computational Biology, Paris, France, September 2003
E. Garcia	Society for General Microbiology, Manchester, UK, September 2003
	Fourth Tularemia International Conference, Bath, UK, September 2003
	Technology and Crime Fighting Conference, San Francisco, CA, September, 2003
	U.S./UK Virulence Workshop, Santa Fe, New Mexico, September 2003
	3rd Animal Genomics Symposium, Research Triangle Park, NC, October 2003
	University of Washington, Seattle, WA, November 2003
	American Academy of Microbiology, Pathogen Colloquium, Key Largo, FL, November 2003
	University of Chile School of Medicine, Santiago, Chile, December 2003
M. Knize	Second International Workshop on Analytical, Chemical and Biological Relevance of Heterocyclic Aromatic Amines, Graz, Austria, May 2003
F. Lightstone	Environmental Mutagen Society (EMS), Miami, FL, May 2003
N. Liu	Gordon Conference, Radiation Oncology, Ventura, CA, February 2003

C. Manohar	1st John Barnabas Memorial Invited Lecture, Postgraduate School For Biological Studies, Ahmednagar, India, December 2003.
	Department of Energy Low-dose Radiation Research Investigators Workshop, October, 2003
S. McCutchen-Maloney	Biomarker Discovery, Validation and Assay Development Meeting, San Francisco, CA, September 2003
	Human Proteome Organization (HUPO) 2nd Annual & International Union of Biochemistry and Molecular Biology (IUBMB) XIX Joint World Congress, Montreal, Canada, October 2003
	Proteome Society Meeting, San Francisco, CA, November 2003
	The First Bay Area DIGE Users Meeting, San Francisco, CA, December 2003
B. Rupp	American Crystallographic Association Meeting Series 30: 85. Cincinnati, OH, July, 2003
	Protein Interactions: Applications for Therapeutic Discovery Through Advances in Protein Design, Expression and Analysis. San Diego, CA, October 2003
	European Structural Biology Forum, Trieste, Italy, December 2003
L. Stubbs	Plenary lecture, Plant and Animal Genomics Meeting, San Diego, CA, January 2003
	Plenary lecture, Livestock Genomics Conference, Houston TX, October 2003
	Computational Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, November, 2003
L. Thompson	Gordon Research Conference on Radiation Oncology, Ventura, CA, January, 2003
	34th Annual Meeting of the Environmental Mutagen Society (EMS), Miami Beach, FL, May 2003

L. Thompson, cont'd	12th International Congress of Radiation Research, Brisbane, Australia, August 2003
	15th Annual Fanconi Anemia Research Fund Scientific Symposium, Houston, TX, October 2003
W. Wilson	National Lab Training Network (NTLN) Course: Molecular Diagnostic Techniques for the Public Health Laboratory
N. A. Yamada	2003 International Congress on Radiation Research, Brisbane, Australia, August 2003

Editorial Boards

J. Albala	*Expert Reviews In Proteomics*
	Expert Reviews In Molecular Diagnostics
R. Balhorn	*Molecular Reproduction and Development*
	International Journal of Biophysics
	Review of Scientific Instruments
A. Christian	*Expert Reviews In Proteomics*
J. Felton	*Mutation Research*
	Environmental and Molecular Mutagenesis
L. Thompson	*Molecular and Cellular Biology*
	Somatic Cell and Molecular Genetics
	DNA Repair

Societies (Major Roles)

M. Coleman	Protein Society Education Working Group
J. Felton	Environmental Mutagen Society, Chair, Nominating Committee
B. Rupp	Sponsor, 2003 UCLA Seaborg Medal Symposium

L. Thompson	Program Committee, 34th Annual Meeting of the Environmental Mutagen Society, Miami Beach, FL, May 2003

Advisory Committees and Grant Review Panels

M. Coleman	Industry/University Cooperative Research Program Biotechnology grant review panel
M. Cosman	Life Sciences/Information Technology study section for the University of California Discovery Grants
K. Dingley	Government representative on the committee of the Genetic and Environmental Toxicology Association of Northern California
J. Felton	Chair, National Institutes of Health/Environmental Protection Agency Agricultural Health Study Advisory Board
	State of California Prop 65 Cancer Advisory Board
	University of California Toxic Substance Research &Training Program (UCTSR&TP) Executive Board
	National Cancer Institute Grants Subcommittee E (PO1 Oversight Panel)
	Chair for one PO1 review
	National Toxicology Program (NTP)/National Institute of Environmental Health Sciences (NIEHS) Report on Carcinogens (Heterocyclic Amines)- Testimony to Board of Scientific Councilors
	Executive Committee, University of California Davis Cancer Center
K. Fidelis	Organizing Committee member, “Critical Assessment of Protein Structure Prediction”,
R. Langlois	International Advisory Council for the Applied Ecology Research Laboratory

L. Radnedge	Grant review study section on Biodefense, National Institute of Allergy and Infectious Diseases (National Institutes of Health)
B. Rupp	Grant Review Panel University of California BioSTAR
	National Institute of Allergy and Infectious Diseases Institutional Review Board Study Section on Botulinum Biophysics
	North Carolina Biotechnology Center, Science and Technology Development Grants Program
	University of California Los Angeles Chemistry and Biology Department Advisory Council
	The Wellcome Trust, UK
	Nominated for U.S. National Committee to the International Union of Crystallography
L. Stubbs	National Institutes of Health, Genome Study Section
	Department of Energy Biological and Environmental Research Advisory Committee
	Mouse Genome Database Advisory Committee
	University of California Davis Cancer Center Internal Advisory Committee
R. Tebbs	Merced College Biotechnology Advisory Committee
L. Thompson	Chair, Review Committee for Canadian National Cancer Institute, Program Project “DNA Damage Responses in Eukaryotes”(P.I. Susan Lees-Miller)
	Review Panel for National Institutes of Health/National Cancer Institute Program Project on Werner Syndrome, University of Washington, Seattle (P.I. Raymond Monatt)

Meetings Organized

L.Stubbs	Conference co-organizer, Mouse Initiatives V: Genomics of Complex Systems in Biomedical Research, Bar Harbor, ME, July 2003

Public Education and Outreach

J. Albala	Sizzlin' Summer Science Keynote Speaker, Livermore, CA, July 2003
	University of the Pacific Family Camp Workshop Presenter, July 2003
	Kern County Teachers, Edward Teller Educational Center, LLNL, July 2003
	BioCom Radio Broadcast: Protein Expression, Summer 2003
	Edward Teller Symposium, LLNL, Sept 2003
	Gender Studies Panel, University of the Pacific, November 2003
	Kern County Teachers, Edward Teller Educational Center, LLNL, December 2003
D. Barsky	Instructor, Advanced Topics in Bioinformatics, California State University, Hayward
A. Christian	School Science Fair Judge
K. Kulp	Sizzlin' Summer Science Keynote Speaker, Livermore, CA, July 2003
	Invited Lecturer, High School Science Teachers Summer School Program, July 2003
	High School Science Fair Project mentor, December 2003
N. Liu	LLNL Open House
C. Manohar	Microarray Collaboration Group Tour, High School students from Logan, Utah, October 2003
	Microarray Collaboration Group Tour, Granada High School AP Biology Class, November 2003
M. Mendelsohn	TOPScience, Christenson Middle School physical science advisor

J. B. Pesavento	Adjunct Professor, Las Positas College
	Judge, Tri-Valley Science and Engineering Fair, April 2003
E. Skowronski	University of California Davis/United States Department of Agriculture Exotic Newcastle Disease monitoring
L. Stubbs	University of California Davis Genetics Graduate Group, student lectures and interaction (throughout 2003).
	Thesis supervisor and mentor for Master's degree, Sha S. Hammond, California State University Hayward, completed 12/03
	LLNL Project supervisor for Master's thesis work, Samir Padurangi, University of California Davis Bioinformatics, 6/03-present.
	Adjunct Professor, University of Tennessee, Knoxville/Oak Ridge National Laboratory, Department of Genome Sciences
W. Wilson	Speaker, Teacher Research Academy, Edward Teller Education Center
N. Yamada	Tri-Valley Science and Engineering Fair, Scientific Review Committee, April 2003

Appendix III: Publications and other information transfer

BBRP utilizes many different methods to transfer information to the public, including those listed in the table below. Data for 2002 is included for comparison. Publication information constitutes the remainder of this appendix.

Information type	CY2002	CY2003
Journal Articles	57	66
Book Chapters	4	4
Patents Issued	7	4
Patents Pending	5	5
Invention Disclosures	6	8
Income from Royalties and Licensing	$260,444	~$190,000
cDNA clones made available	1,253,000	832,000

Manuscripts and book chapters published (2003)

Barton, T. S., Wyrobek, A. J., Hill, F., Robaire, B., and Hales, B. F. (2003). Numerical chromosomal abnormalities in rat epididymal spermatozoa following chronic cyclophosphamide exposure. *Biol. Reprod.*/Epub 69(4), 1150-1157.

Beernink, P. T., Segelke, B. W., and Coleman, M. A. (2003). High-throughput, cell-free protein expression screening using RTS 100. *Biochemica* 1, 4-5.

Brewer, L. R., Friddle, R., Noy, A., Baldwin, E., Martin, S. S., Corzett, M., Balhorn, R., and Baskin, R. J. (2003). Packaging of single DNA molecules by the yeast mitochondrial protein Abf2p. *Biophysical J.* 85, 2519-2524.

Chain, P., Kurtz, S., Ohlebusch, E., and Slezak, T. (2003). An applications-focused review of comparative genomics tools: Capabilities, limitations and future challenges. *Briefings in Bioinformatics* 4(2), 105-123.

Chain, P., Lamerdin, J., Larimer, F., Regala, W., Lao, V., Land, M., Hauser, L., Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D., Hommes, N., Whittaker, M., and Arp, D. (2003). Complete genome sequence of the ammonia oxidizing bacterium and obligate chemolithoautotroph *nitrosomonas europaea*. *J. Bacteriol*. 185(9), 2759-2773.

Coleman, M. A., Miller, K. A., Beernink, P. T., Yoskikawa, D. M., and Albala, J. S. (2003). Identification of chromatin-related protein interactions using protein microarrays. *J. Proteomics* 3(11), 2101-2107.

Colvin, M. E. and Quong, J. N. (2003). DNA-alkylating events associated nitrogen mustard based anticancer drugs and metabolic byproduct acrolein. *Advances in DNA Sequences Specific Agents*: Elsevier, N.Y., pp. 29-46.

Dingley, K. H., Sanchez, E. A., Turteltaub, K. W., Abel, S., Ebeler, S. E., Mitchell, A., Burns, A. E., Steinberg, F. M., and Clifford, A. J. (2003). Effect of dietary constituents with chemopreventive potential on the bioavailability and adduct formation of a low-dose of the heterocyclic amine 2-amino-1- methyl-6-phenyl-imidazo[4,5-*b*]pyridine (PhIP). *Nutr. Cancer* 46(2), 212-221.

Edwards, T. C., Marsh, A., Sanders, C., Albala, J. S., and Prange, C. K. (2003). Automated construction of an open reading frame library from *Sinorhizobium meliloti*. *J. Assoc. Laboratory Automation* 8(3), 44-49.

Felton, J. S., and Wu, R. W. (2003). Ames/Salmonella assay: A bacterial test for mutagens. *Nature Encyclopedia of Life Sciences*, On Line Publishing, 1-4.

Forde, C. E. and McCutchen-Maloney, S. L. (2003). Characterization of transcription factors by mass spectrometry and the role of SELDI-MS. *Mass Spec. Rev*. 21(6), 419-439.

Garcia, E., Elliott, J., Ramanculov, E., Chain, P., Chu, M., and Molineuz, I. J. (2003). The genome sequence of *Yersinia pestis* bacteriophage ØA1122 reveals an intimate history with the coliphage T3 and T7 genomes. *J. Bacteriol*. 2185(17), 5248-5262.

Gardner, S. and Fernandes, M. (2003). New tools for cancer chemotherapy: Computational assistance for tailoring treatments. *Molecular Cancer Therapeutics* 2, 1079-1084.

Gardner, S. N., Kuczmarski, T. A., Vitalis, E.A., and Slezak, T. R. (2003). Limitations of TaqMan PCR for detecting divergent viral pathogens illustrated by hepatitis A, B, C, and E viruses and human immunodeficiency virus. *J. Clin. Microbiol*. 41, 2417-2427.

Gilbert, M., Edwards, T. C., and Albala, J. S. (2003). Protein expression arrays for proteomics. In: Protein Arrays, Humana Press.

Gilbert, M., Edwards, T. C., Prange, C., Malfatti, M., McConnell, I. R., and Albala, J. S. (2003). Miniaturized protein production for proteomics. In: Protein Arrays, Biochips, and Proteomics, the Next Phase of Genomic Discovery. Albala, J.S., and Humphery-Smith, I., eds., Marcel Dekker, New York, NY, 203-216.

Glaser, R. L., Broman, K. W., Schulman, R. L., Eskenazi, B., Wyrobek, A. J., and Jabs, E. W. (2003). The paternal age effect in apert syndrome is due in part to the increased frequency of mutations in sperm. *Am. J. Hum. Genet*./Epub 73(4), 939-947.

Hammond, S. (2003). Analysis of hypothetical promoter domains of DKFZP564A1164, NPHS1 and HSPOX1 genes. Cal State Hayward – Thesis.

Henderson, P. T., Delaney, J. C., Muller, J. G., Neeley, W. L., Tannenbaum, S. R., Burrows, C. J., and Essigmann, J. M. (2003). Determination of the in vivo mutagenic

potential of the two major hydantoin products of 7,8-dihydro-8-oxoguanine oxidation. *Biochemistry* 42, 9257-9262.

Hill, F., Marchetti, F., Liechty, M., Bishop, J., Hozier, J., and Wyrobek, A. J. (2003). A new FISH assay to simultaneously detect structural and numerical chromosomal abnormalities in mouse sperm. *Mol. Reprod Dev*. 66(2), 172-180.

Hvidsten, T. R., Kryshtafovych, A., Komorowski, J., and Fidelis, K. (2003). A novel approach to fold recognition using sequence-derived properties from sets of structurally similar local fragments of proteins. *Briefings in Bioinformatics* 1(1), 1-10.

Iwobi, A., Heeseman, J., Garcia, E., Igwe, E., Noelting, C., and Rakin, A. (2003). Novel virulence-associated type II secretion system unique to high-pathogenicity *Yersinia enterocolitica*. *Infection and Immunity* 71(4), 1872-1879.

Kantardjieff, K., and Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science* 12, 1865-1871.

Kim, J., Kollhoff, A., Bergman, A., and Stubbs, L. (2003). Methylation-sensitive binding factor YY1 to an insulator sequence within the paternally expressed imprinted gene, *Peg*3. *Human Molecular Genetics* 12(3), 233-245.

Knize, M. G., Salmon, C. P., and Felton, J. S. (2003). Mutagenic activity and heterocyclic amine carcinogens in commercial pet foods. *Mutat. Res.* 539,195-201.

Kohane, M., Daugela, A., Kutomi, H., Charlson, L., Wyrobek, A., and Wyrobek, J. (2003). Nanoscale *in vivo* evaluation of the stiffness of *Drosophila melanogaster* integument during development. *J. Biomed. Mat. Res*. 66A(3), 633-642.

Kulp, K. S., Fortson, S. L., Knize, M. G., and Felton, J. S. (2003). An *in vitro* model system to predict the bioaccessibility of heterocyclic amines from a cooked meat matrix. *Food Chem. Toxicol*. 41, 1701-1710.

Kulp, K. S., Knize, M. G., Fowler, N., Salmon, C. P., and Felton, J. S. (2003). PhIP metabolites in human urine after consumption of well-cooked chicken. *J. Chromatography B*, 2003.

Malba, V., Maxwell, R., Evans, L. B., Barnhardt, A. F., Cosman, M., and Yan, K. (2003). Laser-lathe lithography - a novel method for manufacturing nuclear magnetic resonance microcoils. *Biomedical Microdevices* 5(1), 21-27.

Marchetti, F., Bishop, J. B., Cosentino, L., Moore, D., and Wyrobek, A. J. (2003). Paternally transmitted chromosomal aberrations in zygotes determine their embryonic fate. *Biol. Reprod*. 70, 616-624.

Marchettti, F., Knize, M. G., Ciarappa-Zucca, M. L., Pletcher, R. J., and Layton, D. W. (2003). Biodegradation of potential diesel oxygenate additives: dibutyl maleate (DBM), and tripropylene glycol methyl ether (TGME). *Chemosphere* 52(5), 861-868.

McBride, M. T., Gammon, S., Pitesky, M., O'Brien, T. W., Smith, T., Aldrich, J., Langlois, R. G., Colston, B., and Venkateswaren, K. S. (2003). Multiplexed liquid arrays for simultaneous detection of simulants of biological warfare agents. *Anal. Chem*. 75, 1924-1930.

McBride, M. T., Masquelier, D., Hindson, B. J., Makarewicz, A. J., Brown, S., Burris, K., Metz, T., Langlois, R. G., Tsang, K. W., Anderson, D. A., Venkateswaran, K. S., Milanovich, F. P., and Colston Jr., B. W. (2003). Autonomous detection of aerosolized *Bacillus anthracis* and *Yersinia pestis*. *Anal. Chem*. 75, 5293-5299.

Mielke, S. P. and Krishnan, V. V. (2003). Protein structural class identification directly from NMR spectra using averaged chemical shifts. *Bioinformatics* 19(16), 2054-2064.

Miller, K. A., Sawicka, D., Barsky, D., and Albala, J. S. (2003). Domain mapping of the Rad51 paralog protein complexes. *Nucleic Acids Res*. 32(1), 169-178.

Mohrenweiser, H. W. (2003). Pharmacology and toxicology in the genomics era - Harvesting nuggets without drowning in the sluice. Genes, Pills and People: The effect of pharmacogenomics on individuals and groups. In: *Chapter 2 Pharmacogenomics*, John Wiley & Sons: 29-49.

Mohrenweiser, H. W., Wilson, D. M. III, and Jones, I. M. (2003). Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes. *Mutat. Res. Fundamental and Molecular Mechanisms of Mutagenesis*, 526, 93-125.

Morrow, J. A., Lu, B., Hochtl, P., Oberg, K. A., Rupp, B., and Weisgraber, K. H. (2002). Apolipoprotein E4 forms a molten globule: A potential basis for its association with disease. *J. Biochem*. 277(52), 50380-50385.

Murali, N. and V.V. Krishnan (2003). A primer in nuclear magnetic relaxation in liquids. *Concept. Mag. Reson.* 17(A), 86-116.

Ngu-Schwemlein, M., Corzett, M., Thornton, K. H., Balhorn, R., and Cosman, M. (2003). Extracellular domain of myeline oligodendrocyte glycoprotein (MOG) exhibits solvent-dependent conformational transitions. *Protein and Peptide Letters* 10, 483-490.

Ovcharenko, I., and Loots, G. (2003). Finding the needle in the haystack: Computational strategies for discovering regulatory sequences in genomics. *Current Genomics* 4, 557-568.

Perreault, S. D., Aitken, R. J., Gordon Baker, H. W., Everson, D. P., Huszar, G., Irvine, D. S., Morris, I. D., Morris, R. A., Robbins, W. A., Sakkas, D., Spano, M., and Wyrobek, A. J. (2003). Integrating new tests of sperm genetic integrity into semen analysis. *Adv. Exp. Med. Biol*. 518, 253-268.

Peterson, L. A., Vu, C., Hingerty, B. E., Broyde, S., and Cosman, M. (2003). Solution structure of an O^6-[4-oxo-4-(3-pyridl)buty1]guanine adduct in an 11-mer DNA duplex: evidence for formation of a base 'triplex'. *Biochemistry* 42(45), 13134-13144.

Ramirez, L. M., Axelrod, H. L., Herron, S. R., Rupp, B., Allen, J. P., and Kantardjieff, K. (2003). A high resolution crystal structure of ferricytochrome c' from *Rhodobacter sphaeroides*. *J. Chemical Crystallography* 33(5/6), 413-424.

Reddy, V., Swanson, S., Sacchettini, J. C., Kantardjieff, K. A., Segelke, B., and Rupp, B. (2003). Effective electron density map improvement and structure validation on a Linux multi-CPU web cluster: The TB structural genomics consortium bias removal web service. *Acta Crystallographica D Biological Crystallography* 59(12), 2200-2210.

Rothkamm, K., Kruger, I., Thompson, L. H., and Lobrich, M. (2003). Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol*. 23, 5706-7615.

Rupp, B. (2003). High throughput crystallography at affordable cost: The TB structural genomics consortium crystallization facility. *Accounts of Chemical Research* 36, 173-181.

Rupp, B. (2003). Maximum likelihood crystallization. *J. Struct. Biol*. 142, 162-169.

Shannon. M., Hamilton, A. T., Gordon, L., Branscomb, E., and Stubbs, L. (2003). Differential expansion of zinc-finger transcription factor loci in homologous human and mouse gene clusters. *Genome Res*. 13, 1097-1110.

Sigurdson, A. J., and Jones, I. M. (2003). Second cancers after radiotherapy- any evidence for radiation-induced genomic instability? *Oncogene Reviews* 22, 7018-7027.

Slezak, T., Kuczmarski, T., Ott, L., Torres, C., Medeiros, D., Smith, J., Truitt, B., Mulakken, N., Lam, M., Vitalis, E., Zemla, A., Zhou, C. E., and Gardner, S. (2003). Comparative genomics tools applied to bioterrorism defense. *Briefings in Bioinformatics* 4(2), 133-149.

Smith, T. R., Miller, M. S., Lohman, K., Lange, E. M., Case, L. D., Mohrenweiser, H. W., and Hu, J. J. (2003). Polymorphisms of *XRCC1* and *XRCC3* genes and susceptibility to breast cancer. *Cancer Letters* 190, 183-190.

Smith, T. R., Levine, E. A., Perrier, N., Miller, S. M., Freimanis, R. I., Lohman, K., Case. L. D., Xu, J., Mohrenweiser, H. W., and Hu, J. J. (2003). DNA repair genetic

polymorphisms and breast cancer risk. *Cancer Epidemiology, Biomarkers and Prevention* 12, 1200-1204.

Sorensen, K. J., Turteltaub, K., Vrankovich, G., Williams, J., and Christian, A. T. (2003). Whole-genome amplification of DNA from residual cells left by incidental contact. *Analytical Biochemistry* 324, 312-314.

Tebbs R. S., Thompson, L. H., and Cleaver, J. E. (2003). Rescue of *XRCC1* knockout mouse embryo lethality by transgene-complementation. *DNA Repair* 2(12), 1405-1417.

Terwilliger, T. C., and Rupp, B. (2003). The TB structural genomics consortium: A resource for *Mycobacterium tuberculosis* biology. *Tuberculosis* 83, 223-249.

Thompson, L. H. and Limoli, C. L. (2003). Origin, recognition, signaling, and repair of DNA double-strand breaks in mammalian cells. In: Eukaryotic DNA Damage, Surveillance and Repair, K. Caldecot, ed. Eurekah.com, Inc.

Tucker, J. D., Moore, D. H., Ramsey, M. J., Kato, P., Langlois, R. G., Burroughs, B., Long, L., and Garr, V. F. (2003). Multi-endpoint biological monitoring of phosphine workers. *Mutat. Res*. 536, 7-14.

Venclovas, C. (2003). Comparative modeling in CASP5: Progress is evident, but alignment errors remain a significant hindrance. *Proteins: Struct. Function Genet*. 53, 380-388.

Weinstein, B., and Heller, A. (2003). Biology and Biotechnology Timeline and Compilation of Information 1963-2003. *40 years of Discovery* 1, 1-60.

Wilczynski, B., Kryshtafovych, A., Komorowski, J., Hvidsten, T., Stubbs, L., and Fidelis, K. (2003). A rule-based framework for gene regulation pathways discovery. *The IEEE Computer Society Bioinformatics Conference Paper.*

Winthrop, M. D., DeNardo, S. J., Albrecht, H., Mirick, G. R., Kroger, L. A., Lamborn, K. R., Venclovas, C., Colvin, M. E., Burke, P. A., and DeNardo, G. L. (2003). Selection and characterization of anti-MUC1 scFvs intended for targeted therapy. *Clin. Cancer Res.* 9(1) Part 2 Suppl 2, 3845S-3853S.

Wu, R. W., Panteleakos, F. N., and Felton, J. S. (2003). Development and characterization of CHO repair-proficient cell lines for comparative mutagenicity and metabolism of heterocyclic amines from cooked food. *Environ. Mol. Mutagen.* 41(1), 7-13.

Yin, E., Nelson, D. O., Coleman, M. A., Peterson, L. E., and Wyrobek, A. J. (2003). Gene expression changes in mouse brain after exposure to low-dose ionizing radiation. *Int. J. Radiat. Biol*. 79(10), 759-775.

Zemla, A. (2003). LGA - a method for finding 3D similarities in protein structures. *Nucleic Acids Res.* 31(13), 3370-3374.

Manuscripts and book chapters in press (2003)

Chepanoske, C. L., Brown, K., Turteltaub, K. W., and Dingley, K. (2003). Characterization of a peptide adduct formed by N-acetoxy-2-amino-1-methyl-6-phenylimidazo]pyride (PhIP), a reactive intermediate of the food carcinogen PhIP. *Food Chem. Toxicol*.

Dingley, K., Ubick, E., Vogel, J. and Haack, K. W. (2003). DNA isolation and sample preparation for quantitation of adduct levels by Accelerator Mass Spectrometry. Method Mol. Biol.: Protocols in Molecular Toxicology, Humana, Totowa, NJ.

Gardner, S. N. and Fernandes, M. (2003). Cytostatic anticancer drug development. *J. Experimental Therapeutics and Oncology*.

Gardner, S. N. and Fernandes, M. (2003). The Redemption of Sisyphus: Clinical anticancer drug development and the "Single Gate" initiative. *J. Experimental Therapeutics and Oncology*.

Glaser, R. L., Broman, K. W., Schulman, R. L., Eskenazi, B., Wyrobek, A. J. and Jabs, E. W. (2003). The paternal age effect in apert syndrome is due in part to the increased frequency of mutations in sperm. *Am. J. Hum. Genet*.

Hamilton, A. T., Huntley, S., Kim, J., Branscomb, E., and Stubbs, L. (2003). Lineage-specific expansion of KRAB zinc-finger transcription factor genes: implications for the evolution of vertebrate regulatory networks. *Cold Spring Harbor Symposium on Quantitative Biology, Volume 68: The Genome of Homo Sapiens*, 68.

Huntley, S., Hamilton, A. T., Kim, J., Branscomb, E., and Stubbs, L. (2003). Tandem gene family expansion and genomic diversity. *Comparative Genomics: A guide to the analysis of eukaryotic genomes*.

Kantardjieff, K. A., and Rupp, B. (2003). Protein isoelectric point as a predictor for increased crystallization screening efficiency. *Bioinformatics*.

Kim, J., Bergmann, A., Lucas, S., Stone, R., and Stubbs, L. (2003). Lineage-specific imprinting and evolution of the zinc finger gene *ZIM2*. *Genomics*.

Kryshtafovych, A., Hvidsten, T. R., Komoroski, J., and Fidelis, K. (2003). Fold recognition using sequence fingerprints of protein local substructures. *Proceedings of the IEEE Computer Society Bioinformatics Conference.*

Martin, E. A., Brown, K., Gaskell, M., Al-Azzawi, A., Garner, R. C., Boocock, D. J., Mattock, E., Pring, D. W., Dingley, K., Turteltaub, K. W., Smith, L. L., and White, I. N.

H. (2003). Tamoxifen DNA damage detected in human endometrium using the ultra sensitive technique of accelerator mass spectrometry. *Cancer Res*.

Mertens, A. C., Mitby, P. A., Perentesis, J. P., Radloff, G., Kiffmeyer, W. R., Neglia, J. P., Meadows, A., Jones, I., Potter, J. D., Freidman, D., Yasui, Y., Robinson, L. L., and Davies, S. M. (2003). XRCC1 and glutathione S-transferase polymorphism and susceptibility to therapy-related cancer in Hodgkin's disease survivors: A report from the Childhood Cancer Survivor Study. *Clinical Cancer Research.*

Mohrenweiser, H. W. (2003). Genetic variation and exposure related risk estimation: Will toxicology enter a new era? DNA repair and cancer as a paradigm. *J. Toxicologic Pathology*.

Ovcharenko, I., Loots, G. G., Bofelli, D., and Rubin, E. M. (20030). eShadow: a tool for comparing closely related sequences. *Genome Res.*

Ovcharenko, I., Loots, G. G., Hardison, R., Miller, W., and Stubbs, L. (2003). *zPicture*: Dynamic alignment and visualization tool for analyzing conversation profiles. *Genome Res*.

Radnedge, L., Agron, P. G., Hill, K. K., Jackson, P. J., Ticknor, L. O., Keim, P., and Andersen, G. L. (2003). Genome differences that distinguish *B. anthracis* from closely related *Bacillus cereus* and *Bacillus thuringiensis*. *App. Environ. Microbial.*

Segelke, B. W., Schafer, J., Coleman, M. A., Lekin, T. P., Toppani, D., Skowronek, K. J., Kantardjieff, K. A., and Rupp, B. (2003). Laboratory scale structural genomics. *J. Struct. Func. Genomics*.

Slezak, T. R. and Salzberg, S. L. (2003). Bioinformatics methods for microbial detection and forensic diagnostic design. *Microb. Forensics* (ed. Bruce Budlowle, Steven Schutzer and Roger Breeze) Elsevier Science.

Sloter, E., Nath, J., Eskanazi, B., and Wyrobek, A. J. (2003). Effects of male age on the frequencies of germinal and heritable chromosomal abnormalities in humans and rodents: a review of the literature. *Fertil. Steril.*

Xi, T., Jones, I. M., and Mohrenweiser, H. W. (2003). Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. *Genomics*.

Appendix IV: Patents and Invention Disclosures

Patents Issued – 2003		
Inventors	Title	Issue Date
M. Allen Northrup	Microfabricated Sleeve Devices for Chemical Reactions	2/25/2003
M. Allen Northrup, Yun-Tai Hsueh, Rosemary L. Smith	Microfabricated Electrochemiluminescence Cell for Chemical Reaction Detection	2/18/2003
Stefan P. Swierkowski	Injector-Concentrator Electrodes for Microchannel Electrophoresis	5/6/2003
Ramakrishna S. Madabhushi, Stuart A. Gammon	Polyacrylamide Medium for the Electrophoretic Separation of Biomolecules	11/11/2003

Patents Pending – 2003		
Inventors	Title	Tracking No.
Krzysztof A. Fidelis, Andriy A. Kryshtafovych	Local Descriptors of Protein Structure	10728
Shea Nicole Gardner	Modeling Multi-Drug Chemotherapy: Tumor Control, Resistance, and Toxicity	10955
Allen T. Christian, Brent Segelke, Bernard Rupp	Augmented Automated Macromolecular Crystal Detection from Light Microscopy Images	11013
Allen T. Christian, Janelle S. Lamberton	Gene Silencing Using DNA:RNA-Hybrid Short, Interfering Molecules	11065
Andrew J. Wyrobek, Francesca S. Hill, Francesco O. Marchetti	Multilabeling Mouse FISH Assay for Detecting Structural and Numerical Chromosomal Abnormalities	11097

Appendix IV: Patents and Invention Disclosures, cont'd

Invention Disclosures - 2003		
Inventors	Title	Tracking No.
Jeffrey Stuart Kallman	Aerogel Thin Film Frustrated Total Internal Reflection Imaging Acoustic Sensor	11142
Peter T. Beernink, Matthew A. Coleman, Brent W. Segelke	High-Throughput Protein Production Screening	11157
Adam Tadeusz Zemla	Local-Global Alignment - A Method For Finding 3D Similarities In Protein Structures	11160
Paul T. Henderson	Light-Mediated Electrochemical DNA Detection	11212
Andrew J. Wyrobek, Christine Hartmann-Siantar, Francesco Marchetti	Rapid Non-Invasive Indicators of Tissue-Specific Cellular Damage	11223
Brian E. Souza, Matthew A. Coleman, James W. George, Sephan P. Velsko	Novel Epitope Identification by Antibody Specific Immunopanning	11227
Rodney L. Balhorn	Antibody Mimics, Selective High Affinity Ligands for Malignant Diseases - Diagnosis and Treatment	11262
Rodney Balhorn, Monique Cosman, Ramakrishna S. Madabhushi	Application of Light Emitting Polymers to Detect DNA Adducts in Blood	11290

Appendix V: Funding information (listed by sponsoring agency)

Work For Others:

Sponsor	*Support (K$)*	*Principal Investigator*	*Objective*
ARMY	845.2	K. Dingley	
HHS/CDCP	482.2	K. Fidelis	To support the continuing operation of the Critical Assessment of Protein Structure Prediction (CASP) process and to expand its infrastructure,
HHS/CNRR	2,533.7	K. Turteltaub	Develop and demonstrate biomedical applications of AMS
HHS/NCI	43.8	I. Jones	Measure the level of DNA damage using an assay developed in the BBR Program
HHS/NCI	293.7	J. Albala	To elucidate the relationship of a newly identified human gene, RAD51B, to breast cancer-related genes
HHS/NCI	2,524.9	J. Felton	Determine whether heterocyclic amine mutagens/carcinogens contribute to human cancer incidence
HHS/NCI	245.0	K. Kulp	Develop an exposure assessment method that will predict the effect of digestion parameters, intestinal trasport and diet upon the low dose exposure to commonly ingested food carcinogens
HHS/NCCRM	280.2	K. Kulp	Studying the interactions of Flor-Essence^{R} Tonic & PhIP

HHS/NCI	364.7	L. Thompson	To understand the molecular regulatory processes cells use to minimize genetic damage and genetic instability associated with reactive oxygen species arising from endogenous processes or ionizing radiation
HHS/NCI	324.0	N. Liu	To understand the role of XRCC2 in DNA repair, specifically in homologous recombinational repair pathways
HHS/NIAID	528.5	E. Garcia	Gain understanding as to how acute and highly lethal bacterial pathogen such as Yersinia pestis has evolved.
HHS/NCI	1,000.0	C. Prange	Isolate large numbers of full-length genes and make them publicly available.
Los Alamos National Lab	632.5	B. Rupp	To carry out large scale crystallizations of proteins from Tuberculosis
OFA	255.0	A. Christian	Identification of human population groups
Zyomyx CRADA	125.5	J. Albala	A miniaturized system for protein production in Baculovirus.
UC Berkeley	138.7	A. Wyrobek	Investigate the relationship between aneuploidy sperm and aneuploidy at birth caused by parental age, diet, and smoking

UC Berkeley	268.3	K. Turteltaub	Address the effects of chemical dose on the absorption and the associated DNA and protein adduct levels for exposure to benzene and trichlorethylene, carcinogens found at superfund waste sites and common to the urban environment.
UC Breast Cancer Research	162.0	K. Kulp	Evaluation of essiac tea to prevent mammary tumors and improve our understanding of the impact of a complementary and alternative therapy
UC/Biostar	83.4	J. Albala	A miniaturized system for protein production in Baculovirus.
TOTAL	**11,131.2**		

DOE (Office of Biological & Environmental Research):

Support (K$)	***Principal Investigator***	***Objective***
166.0	M. Colvin	Funds 25% of PI for assignment to OBER to assist with new programs in computational biology
206.0	G. Andersen	Understand protein function by understanding what proteins in the microbial cell specifically bind to one another
913.0	L. Thompson	Understand the relative contributions of the individual DNA-damage response pathways to the recovery of mammalian cells from exposure to ionizing radiation (IR) in the range of 0-1 Gy
161.0	M. Coleman	Build a computer-based gene-network model of pathways involved in cellular IR-response, understand the genomic regulation of radiation effects on transcription, identify susceptibility factors, and predict new genes that may be associated with low-dose IR exposure
728.0	A. Wyrobek	Provide mechanistic and molecular knowledge of the cellular response to low dose IR to help reduce the uncertainty in assessing health risks at low-dose levels, and identify candidate genes responsible for differential susceptibility to low-dose IR exposure

6,000.0	JGI	Make significant contribution to the international Human Genome Project
200.0	C. Prange	The I.M.A.G.E. Consortium will work in conjunction with various external groups to isolate large numbers of cDNA clones
400.0	L. Stubbs	Development of a mouse mutant resource to link human genes to health-related functions
500.0	L. Stubbs	Characterization of a region of HSA19q13.4 containing imprinted genes, and study the evolutionary history by comparative analysis of genomic sequence from multiple species
1,620.0	L. Stubbs	Functional annotation of genes and regulatory sequences on HSA19
64.0	K. Fidelis	Advance the computational approaches to the structural and functional studies to genomic sequences
58.0	L. Ashworth	Support for scientific input into the design, development and implementation of both the materials development and educational modules of the Einstein Institute for Science Health and the Courts (EINSHAC) program

225.0	K. Turteltaub	Develop unique methodology for the analysis of a long-lived isotope of beryllium, beryllium-10, to apply it to understanding beryllium dosimetry and to help determine the cellular and molecular mechanisms responsible for beryllium disease
11,241.0	**TOTAL**	

DOE (NNSA and DHS):

Support (K$)	***Principal Investigator***	***Objective***
419.6	R. Balhorn	Toxins
318.0	G. Andersen	Microbial Background
1,852.5	A. Christian	BioBriefcase
50.0	E. Garcia	BDAP
503.0	E. Garcia	Profiling Transcriptional Response
2,600.0	P. McCready	BASIS
834.2	P. McCready	Human Assay Development
945.8	P. McCready	Tailored Assays for Detection of Agro Terrorism
500.0	P. McCready	ABSAT
824.5	P. McCready	Viral Genomics
850.0	S. McCutchen-Maloney	Pathogen Pathways
1,346.1	T. Slezak	BioInformatics
11,449.3	**TOTAL**	

LLNL (Laboratory Directed Research and Development):

Support (K$)	*Principal Investigator*	*Objective*
190.3	P. Beernink	Assess the effects of mutations that cause human diseases using computational and experimental approaches
190.0	A. Zemla	Develop a highly effective method of characterizing newly dequenced proteins in terms of their 3-D structure and function
189.0	H. Mohrenweiser	Tools for quantitative studies of the impact of complex patterns of genetic variation on cellular processes
1,149.4	P. Fitch	Identify and characterize genes and proteins responsible for virulence and develop a model of the virulence mechanisms of specific bacterial pathogens
283.0	M. Colvin	Strategic Initiative in Applied Biological Simulations
1,983.3	K. Turteltaub	Develop an approach that will rapidly lead to assays to detect protein and metabolite signatures that will identify a BW threat agent in a variety of situations
30.0	M. Thelen	Letant Collaboration
125.0	L. Stubbs	Quong Collaboration
301.1	A. Christian	Develop and apply technical capabilities that will ultimately create a new LLNL core competency to measure, manipulate and predictively model life at the level of individual cells.

330.0	M. Thelen	Develop novel biological and computational technologies that will be used to characterize signature biochemical pathways in selected microbes
802.2	K. Turteltaub	Develop an approach that will rapidly lead to assays to detect protein and metabolite signatures that will identify a BW threat agent in a variety of situations
36.0	M. Corzett	Hollars Collaboration
128.6	R. Balhorn	Develop the methodology needed to synthesize High Affinity Ligands (HALs) - small organic molecules that function like synthetic antibodies and bind to specific biological targets with high affinity and specificity
116.0	K. Fidelis	Develop a new approach to identification of gene regulation mechanisms based on a supervised machine learning method.
100.0	K. Fidelis	Develop a global database of theoretically derived protein structures.
5,953.9	**TOTAL**	

LLNL (General and Administrative):

Support (K$)	***Principal Investigator***	***Objective***
187.0	M. Mendelsohn	Melanoma Investigation
187.0	**TOTAL**	